(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
8 January 2004 (08.01.2004)

PCT

(10) International Publication Number
WO 2004/003204 A1

(51) International Patent Classification[7]: C12N 15/31

(21) International Application Number: PCT/KR2003/001279

(22) International Filing Date: 28 June 2003 (28.06.2003)

(25) Filing Language: Korean

(26) Publication Language: English

(30) Priority Data:
10-2002-0037718 29 June 2002 (29.06.2002) KR

(71) Applicants *(for all designated States except US)*: **KOREA REASERCH INSTITUTE OF BIOSCIENCE AND BIOTECHNOLOGY** [KR/KR]; 52 Eoeun-dong, Yuseong-gu, Daejeon 305-333 (KR). **LEADBIO, INC** [KR/KR]; #1721 Hanjin Resort Officetel, 535-5 Bongmyong-dong, Yuseong-gu, Daejeon 305-333 (KR). **BIOHOLDINGS CO., LTD** [KR/KR]; #201 BVC/KRIBB, 52 Eoeun-dong, Yuseong-gu, Daejeon 305-333 (KR).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **KANG, Hyun-Ah** [KR/KR]; #102-1202 HANA Apt., Sinseong-dong, Yuseong-gu, Daejeon 305-345 (KR). **RHEE, Sang-Ki** [KR/KR]; #102-1202 HANA Apt., Sinseong-dong, Yuseong-gu, Daejeon 305-345 (KR). **SOHN, Min-Jeong** [KR/KR]; 175-15 Deokam-dong, Daedeok-gu, Daejeon 306-200 (KR). **KIM, Jeong-Yoon** [KR/KR]; #102-1202 HANA Apt., Sinseong-dong, Yuseong-gu, Daejeon 305-345 (KR).

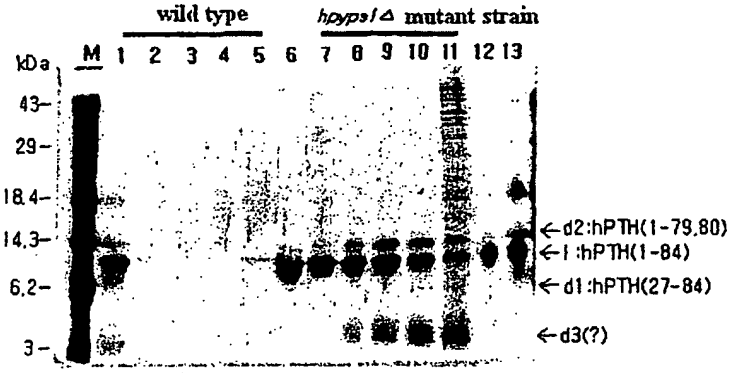
(74) Agent: **SON, Min**; 17th Floor, City Air Tower, 159-9 Samsung-dong, Gangnam-gu, Seoul 135-973 (KR).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),

*[Continued on next page]*

(54) Title: HANSENULA POLYMORPHA YAPSIN DEFICIENT MUTANT STRAIN AND PROCESS FOR THE PREPARATION OF RECOMBINANT PROTEINS USING THE SAME




**Lane M: Molecular marker**

**1: wild type, o-h reaction**
**2: wild type, 2-h reaction**
**3: wild type, 4-h reaction**
**4: wild type, 6-h reaction**
**5: wild type, 24-h reaction**
**6: distilled water + hPTH, o-h reaction**
**7: mutant strain, o-h reaction**
**8: mutant strain, 2-h reaction**
**9: mutant strain, 4-h reaction**
**10: mutant strain, 6-h reaction**
**11: mutant strain, 24-h reaction**
**12: hPTH 100 ng**
**13: hPTH 200 ng**

(57) Abstract: The present invention relates to polynucleotide containing *HpYPS1* gene coding *Hansenula polymorpha* yapsin1; polypeptide coded thereby; *Hansenula polymorpha* mutant wherein yapsin activity is reduced by the mutation of the *HpYPS1* gene coding *Hansenula polymorpha* yapsin1; recombinant *Hansenula polymorpha* strain expressing a foreign protein prepared by introducing a gene coding a foreign protein to the *Hansenula polymorpha* mutant; and a method for preparing a foreign protein comprising the steps of culturing the recombinant*Hansenula polymorpha* strain under the condition that the foreign protein can be expressed, and isolating the foreign protein from the obtained culture broth.

WO 2004/003204 A1

Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

**Published:**
— *with international search report*

# *HANSENULA POLYMORPHA* YAPSIN DEFICIENT MUTANT STRAIN AND PROCESS FOR THE PREPARATION OF RECOMBINANT PROTEINS USING THE SAME

## Technical Field

The present invention relates to a process for producing a recombinant protein secreted and expressed from methanol-utilizing *Hansenula polymorpha* yeast at a high yield by efficiently preventing decomposition of the recombinant protein.

More preferably, it relates to a process for efficiently producing a recombinant protein by destroying yapsin1 gene of *H. polymorpha* strain to prevent decomposition of the protein containing a basic or dibasic amino acid residue produced in *H. polymorpha*.

More particularly, it relates to a process for producing a recombinant protein in an intact configuration at a high yield in *H, polymorpha* by preventing decomposition of the recombinant protein having a basic or dibasic amino acid residue in the protein, such as human parathyroid hormone, human serum albumin and serum albumin fusion protein, which comprises cloning the *HpYPS1* gene encoding *Hansenula polymorpha* aspartic protease type yapsin1, preparing a *H. polymorpha* strain with the defective *HpYPS1* gene using the cloned gene and culturing a transformant transformed from the strain as a host.

**Background Art**

Recently, as demand for high-purity protein medicaments is suddenly increased due to increase of incurable diseases and improvement of public medical standard, relative importance of medicinal recombinant proteins in the health-related bioengineering field is highly raised.

Therefore, the frequency of use of yeast which is a monocellular eukaryotic microorganism as a host system for mass-production of a recombinant protein is gradually increased. Particularly, since yeast has the protein secretion route very similar to those of higher animal cells, it is habitually used as a microorganism host system for production of human-derived secretion proteins. Also, since most kinds of yeast normally secret a very small number of proteins out of the cell, advantageously, recombinant proteins secreted from yeast can be readily recovered and purified. In recent, mass-production of serum proteins, vaccines and other various important medical proteins using non-traditional yeasts including *Hansenula polymorpha* and *Pichia pastoris*, other than the traditional yeast *Saccharomyces cerevisiae* has been successively conducted (Gellissen G., Appl. Microbiol. Biotechnol. 54, 741 (2000)).

As a eukaryotic microorganism, yeast secretes proteins by the substantially same method with mammal cells and involves similar protein modification and cleavage procedures. A protein which has undergone the secretion route becomes to have its final 3-dimensional structure at the Endoplasmic reticulum. In case of glycoprotein, N- and O-

bonding sugar chains are attached thereto. Subsequently, the protein is transferred to the Golgi apparatus, in which it is further subjected to the protein modification procedures such as trimming of oligosaccharide or protein cleavage, and thereafter, is transferred to different organs, inserted into components of the cell membrane or secreted out of the cell.

As described above, since the protein secretion procedures in yeast involve various kinds of post-translational modification processes, the secretion and production of a foreign protein in yeast may cause many problems. Particularly, when a recombinant protein is secreted and produced in yeast, it is necessary to use an efficient expression and secretion system in order to increase productivity, but is also important to prevent decomposition of the produced and secreted foreign protein. If a recombinant yeast is cultured for a long period of time at a high concentration in a fermenter, proteases which are naturally secreted from the host cell or exist in the cell through cell lysis are released to medium and degrade the produced recombinant proteins, thereby causing reduction in overall productivity of the recombinant proteins. In order to solve this problem, for yeasts including *Saccharomyces cerevisiae*, *Hansemula polymorpha*, *Pichia pastoris* and the like, which have been used as recombinant protein expression systems, various protease deficient strains have been developed. Primarily, strains, in which *PEP4*, *PRB1*, or *CPY* genes encoding degradative enzymes existing in yeast vacuole are destroyed (Alvarez et al., J. Biotechnol. 38, 81 (1994); Chen et al., Curr. Genet. 27, 201 (1995); Gleeson et al., Methods Mol. Biol. 103, 81 (1998); Kang et al. In *Hansemula polymorpha* (ed. Gellissen G) p.124 (2001)), have been developed. In addition to the vacuole

degradative enzymes, *kex1*△ strain have been developed, in which *KEX1* gene encoding carboxypeptidase α existing in the Golgi apparatus is destroyed. By the *kex1*△ strain, C-end decomposition of hirudin in *Saccharomyces cerevisiae* (Hinnen et al., In Gene expression in recombinant microorganisms (ed. Smith A.), p121 (1995)), of human epidermal growth factor in *Hansenula polymorpha* (Heo et al., Protein expr. purif. 24, 117 (2001)) and of rodent or human endostatin in *Pichia pastoris* (Boehm et al., Yeast 15, 563-567 (1999)) can be significantly reduced.

Recently, yeast aspartic protease type yapsins having activity to recognize and cut basic amino acids existing as a single or a pair in *Saccharomyces cerevisiae* have been identified, which are novel proteases existing in the cell membrane (Egel-Mitani et al., Yeast 6, 127-137 (1990)). Yapsin1 (also previously known as yeast aspartic protease 3(YAP3)) was firstly known to the public among the yeast aspartic proteases, and yapsin2 (also previously known as MKC7) was known thereafter (Komano and Fuller, Proc. Natl. Acad. Sci. USA 7, 92,10752-10756 (1995)). By the *Saccharomyces cerevisiae* genome information which has been recently disclosed to the public, additional genes encoding at least 5 yapsin type protease presumed to have similar functions, such as yapsin3, yapsin6 and yapsin7, have been reported to exist so far (Olsen et al. Biochem. J. 339, 407-411 (1999)). Though the physiological functions of these yapsins are not clearly shown, as the number of study cases reporting that target recombinant proteins which are intended to secret and produce in *S. cerevisiae* are cleaved by the protease activity of yapsin is increased, yapsin deficient yeast strains attract public attention as an

useful strain for production of a recombinant protein, particularly a foreign peptide having a basic amino acid. Recombinant proteins which have been reported to have problems of being cleaved by yapsins in secretion and production in *S. cerevisiae*, till now, include human serum albumin (Kerry-williams et al., Yeast 14, 161-169 (1998)), human parathyroid hormone (Kang et al., Appl Microgiol Biotechnol., 50, 187-192 (1998)); Korean Patent Registration No. 0246932 (publicated on December 8, 1999)), insect diuretic hormone (Copley et al., Biochem J., 330, 1333-1340 (1998)), glucagon and glucagon-like peptide (Egel-Mitani et al., Enzyme Microb Technol. 26, 671-677 (2000): USA PAT. NO. 6,110,703) and human elafin precursor (Bourbonnais et al., Protein Exp. Purif. 20, 485 (2000)). Meanwhile, considering that *YPS1* deficient *S. cerevisiae* strain shows a considerable progress in decomposition of hPTH at the last stage of the cultivation using a fermenter, the present inventors have developed *S. cerevisiae* yapsin multiple deficient mutant strain (*yps1*$\Delta$/*yps2*$\Delta$/*yps3*$\Delta$), in which the *YPS2* and *YPS3* genes coding for yapsin2 and yapsin3 are removed. As a result, we have obtained an excellent result of preventing 90% or more of degradation of human parathyroid hormone observed in a high-concentration cultivation (Korean Patent Application No. 2000-51267 and International Application No. PCT/KR01/01447).

*H. polymorpha* , one of methanol-utilizing yeasts, is in the spotlight as a very useful yeast host for mass production of recombinant proteins since it has advantages in that strong and controllable promoters are developed, alike *Pichia pastoris,* and a foreign gene can be multiply introduced into the host chromosome (Faber et al., Yeast 11,

1331(1995)). Up to date, various kinds of foreign proteins have been expressed and the expression levels often reached over 1 g/L in case of high-concentration cultivation using a fermenter. Particularly, it has been reported that when recombinant phytase is secreted and produced, the expression level is about 13.5 g/L (Mayer et al., Biotechnol. Bioeng. 63, 373-381 (1999)). Therefore, the *H. polymorpha* expression system becomes distinguished as one of the most potential systems among several presently available eukaryotic cell expression systems. Especially, since some of the recombinant proteins which have been produced in the initial stage in *H. polymorpha* have already passed clinical trials and are on the market (ex., hepatitis B vaccine) or in the product development phase(ex., hirudin), *H. polymorpha* is considered as a suitable expression system for production of a recombinant protein to be developed as an medicament (Gellissen G., Appl Microbiol Biotechnol. 54 741-750 (2000)). Also, as recently getting into the post-genome era, there is an increased need for a high-efficiency expression system for functional analysis of novel genes, and thus it is expected that an expression system using *H. polymorpha* would bear a great part in functional and structural analysis of novel proteins as well as mass production of useful proteins derived from higher eukaryotic cells.

**Disclosure of Invention**

Therefore, the present invention has been made to develop a high-efficiency *H.*

*polymorpha* expression system by solving the problems related to undesired cleavage and decomposition of a recombinant protein by an yapsin type protease, and it is an object of the present invention to provide a technology for secreting and producing a recombinant protein in an intact configuration at a high efficiency by preventing cleavage and decomposition of the recombinant protein expressed in *H. polymorpha*, in which the *HpYPS1* gene encoding *H. polymorpha* yapsin1 is cloned, the cloned gene is used to prepare *HpYPS1* gene deficient mutant strain and the prepared mutant strain is used as a recombinant protein expression host.

In accordance with an aspect, the present invention provides a gene sequence encoding yapsin1(*HpYPS1*) of *H. polymorpha* strain by probing yapsin protease genes in *Hansenula polymorpha* strain, as identified in *S. cerevisiae*.

Also, in further aspect, the present invention provides a yapsin1 polypeptide having *Hansenula polymorpha* strain-derived aspartic protease activity.

Also, in another aspect, the present invention provides a secretion signal gene sequence and a peptide sequence of *HpYPS1* polypeptide for secretion of a foreign protein recombinantly produced in *Hansenula polymorpha* strain.

According to yet another aspect, in order to develop a *Hansenula polymorpha* strain capable of secreting and producing a recombinant protein in an original configuration at a high efficiency, the present invention provides an expression system for secretion and production of a recombinant protein at a high efficiency in *Hansenula polymorpha* by cloning the above-described *Hansenula polymorpha* yapsin protease gene

and deleting the cloned gene to minimize the decomposition of the recombinant protein in *Hansenula polymorpha*.

In order to accomplish the above objects, the present inventors have developed a *Hansenula polymorpha* mutant strain (*hpyps1*△), in which the *HpYPS1* gene is deleted, by cloning the *HpYPS1* gene encoding *Hansenula polymorpha* yapsin1 and subjected the resulting *HpYPS1* gene to a functional analysis. Consequently, we have formed a method for increasing productivity of a recombinant protein by using the *Hansenula polymorpha* mutant strain as a host strain to express the protein containing a basic or dibasic amino acid residue in the protein, including human parathyroid hormone, human serum albumin and human serum albumin fusion protein, thereby significantly reducing the decomposition of the recombinant protein expressed and secreted in *H. polymorpha*.

**Brief Description of Drawings**

Further objects and advantages of the invention can be more fully understood from the following detailed description taken in conjunction with the accompanying drawings in which:

Fig. 1 shows the nucleotide sequence of gene *HpYPS1* encoding *H.polymorpha* yapsin1 and the expected amino acid sequence (①: a signal sequence, ②: the expected GPI anchor, ③: a hydrophobic amino acid-golgi membrane fixing domain);

Fig. 2 shows the comparison of the amino acid sequence between *H.polymorpha*

yapsin1 and *S. cerevisiae* yapsin type proteases;

Fig. 3 is the result of the functional complementation experiment showing that temperature sensitivity of *Saccharomyces cerevisiae* yapsin multiple deficient mutant strain is recovered as the *Hansenula polymorpha* yapsin1 gene *HpYPS1* is expressed;

Fig. 4 is a schematic view showing the procedures for preparing a *Hansenula polymorpha* yapsin1 gene -disrupted mutant strain *hpyps1*$\Delta$;

Fig. 5 shows the results of comparison of protease activity in the culture supernatants of the *H. polymorpha* wild type and the yapsin1-deficient *hpyps1*$\Delta$mutant strain, in which human parathyroid hormone is used as a substrate;

Fig. 6 shows a schematic view of the procedures to construct the human parathyroid hormone expression vector pMOXhPTH for *H. polymorpha*;

Fig. 7A shows the result of Southern blot to examine the expression vector insertion sites and the number of copies by isolating chromosomes of the *Hansenula polymorpha* wild type and the mutant strain *hpyps1*$\Delta$ transformed with the vector pMOXhPTH;

Fig. 7B shows the results of SDS-polyacrylamide electrophoresis of yeast culture supernatants, followed by staining, to compare the expression and decomposition aspects of recombinant human parathyroid hormone in the transformants of *H.polymorpha* wild type and *hpyps1*$\Delta$ mutant strain;

Fig. 8A shows the results of comparison of the expression and decomposition aspects of recombinant human serum albumin when the *H. polymorpha* wild type strain

and the *HpYPS1* gene-deleted *hpyps1*Δ mutant strain are used as a host, in which yeast culture supernatants are analyzed by SDS-polyacrylamide electrophoresis, followed by staining;

Fig. 8B shows the results of comparison of the expression and decomposition aspects of recombinant human serum albumin when the *H. polymorpha* wild type strain and the *HpYPS1* gene-deleted *hpyps1*Δmutant strain are used as a host, in which yeast culture supernatants are analyzed by Western blotting;

Fig. 9 shows a schematic view of the procedures for constructing the albumin-TIMP2 fusion protein expression vector pYHSA13-T2 for *Hansemula polymorpha*; and

Fig. 10 shows the results of the comparison of the expression and decgradation aspects of recombinant albumin-TIMP2 fusion protein when the *H.polymorpha* wild type strain and the *HpYPS1* gene-deleted *hpyps1*Δ mutant strain are used as a host, in which yeast culture supernatants are analyzed by Western blotting. The Lane 1 represents the wild type transformant transformed with pYHSA12(+), the Lanes 2 to 5 represent the wild type transformant transformed with pYHSA13-TIMP2, the Lanes 6 to 9 represent the *hpyps1*Δmutant strain transformed with pYHSA13-TIMP2 and the Lane 10 represent isolated and purified albumin (200 ng).

**Best Mode for Carrying Out the Invention**

*H. polymorpha* is a class of yeast which can use methanol as a carbon source

and energy source. According to the present invention, there is provided a *H. polymorpha* mutant strain with yapsin enzyme destroyed, based on the fact that when *H. polymorpha* is used in a method for producing a foreign protein, the yield of the foreign protein is reduced since the produced foreign protein is decomposed by yapsin enzyme which is an aspartic protease existing in the yeast.

As described herein, the yapsin enzyme family, which is a sub-family of aspartic proteases, has an ability to specifically cleave a basic or dibasic amino acid residue of a protein. This is a difference from other aspartic proteases cleaving hydrophobic residues.

The present inventors have obtained a segment showing similarity with the *S. cerevisiae* yapsin1 encoding gene (*YPS1*) from the PCR-amplified chromosomes derived from *H. polymorpha* and conducted researches and studies on functions of the segment.

As a result, the segment was identified as *HpYPS1* gene of *H. polymorpha*. Further, we have prepared a *Hansenula polymorpha* mutant strain having the *HpYPS1* gene destroyed and confirmed that a foreign protein can be produced from the strain at a high yield. Thus, the present invention has been completed.

In one aspect, the present invention provides a nucleic acid molecule comprising the sequence shown in Fig. 1.

More particularly, the present invention provides a nucleic acid molecule comprising the sequence encoding *Hansenula polymorpha* yapsin1 shown in Fig. 1.

In another aspect, the present invention provides a polypeptide comprising the amino acid sequence shown in Fig. 1.

More particularly, the present invention provides a polypeptide comprising the amino acid sequence of *Hansenula polymorpha* yapsin1(HpYPS1) shown in Fig. 1.

In another aspect, the present invention provides a secretion signal gene sequence and a peptide sequence of HpYPS1 polypeptide for secretion of a foreign protein recombinantly produced in *Hansenula polymorpha* strain, shown as ① in Fig. 1.

In another aspect, the present invention provides a *H. polymorpha* mutant strain having reduced yapsin activity by mutation of the *HpYPS1* gene encoding *H. polymorpha* yapsin1.

In another aspect, the present invention provides a recombinant *H.polymorpha* strain expressing a foreign protein by introducing a gene encoding the foreign protein into the *H. polymorpha* mutant strain.

In another aspect, the present invention provides a process for preparing a foreign protein comprising culturing a recombinant *H. polymorpha* strain under conditions to express the foreign protein and isolating the foreign protein from the culture.

More particularly, the present invention provides a process for secreting and producing a recombinant protein in *Hansenula polymorpha* using a yeast strain lacking the *HpYPS1* gene for a yapsin type protease as a host. The present invention comprises the steps of: cloning gene the *HpYPS1* encoding *H. polymorpha* protease yapsin1; functional analyzing the resulting *H. polymorpha HpYPS1* gene by temperature sensitivity complementation experiment of a *Saccharomyces cerevisiae* yapsin multiple deficient mutant strain; preparing a *hpyps1*Δ mutant strain with the *HpYPS1* gene

destroyed in *Hansenula polymorpha*, followed by analysis of yapsin activity; and preparing a recombinant *Hansenula polymorpha* strain *hpyps1* Δ-pMOXhPTH to express and secrete human parathyroid hormone, a recombinant *Hansenula polymorpha* strain *hpyps1* Δ-pYHSA12 to express and secrete human serum albumin, or recombinant a *Hansenula polymorpha* strain *hpyps1* Δ-pYHSA13-TIMP2 to express and secrete a TIMP2 protein fused with human serum albumin recombinant protein using the *H .polymorpha* mutant strain *hpyps1* Δ as a host, followed by analysis of decomposition level of the recombinant protein. Therefore, the present invention relates to a method for efficiently producing a recombinant protein by preventing degradation of the recombinant protein expressed from methanol utilizing yeast *H. polymorpha*, which comprises cloning the *HpYPS1* gene encoding *H. polymorpha* aspartic protease type yapsin, preparing a *H. polymorpha* strain with the *HpYPS1* gene deficient using the cloned gene, and culturing a transformant transformed from the strain as a host to minimize degradation of the recombinant protein, thereby producing the recombinant protein in an intact configuration at a high yield.

The sequence of the *H. polymorpha YPS1* gene (*HpYPS1*) (SEQ ID NO: 1) cloned according to the present invention was deposited in GenBank under Accession No. AF493990. Also, the *HpYPS1* gene encoding *H. polymorpha* protease yapsin1 was deposited in an international depository authority (Korean Collection for Type Cultures) on the date of June 18, 2002 and assigned Accession No. KCTC 10285BP. Further, the *hpyps1* Δ mutant strain having the *HpYPS1* gene destroyed in *H. polymorpha* was

deposited in an international depository authority (Korean Collection for Type Cultures) on the date of June 18, 2002 and assigned Accession No. KCTC 10281BP. In an embodiment, the recombinant *H. polymorpha* strain *hpyps1* Δ-pMOXhPTH expressing and secreting human parathyroid hormone was deposited in an international depository authority (Korean Collection for Type Cultures) on the date of June 18, 2002 and assigned Accession No. KCTC 10282BP. In another embodiment, the recombinant *H. polymorpha* strain *hpyps1* Δ-pYHSA12 expressing and secreting human serum albumin was also deposited in an international depository authority (Korean Collection for Type Cultures) on the date of June 18, 2002 and assigned Accession No. KCTC 10283BP. In addition, the recombinant *H. polymorpha* strain hpyps1 Δ-pYHSA13-T2 expressing and secreting albumin-TIMP2 fusion protein was deposited in an international depository authority on the date of June 23, 2003 and assigned Accession No. KCTC 10485BP.

Now, the present invention will be explained in detail by the following examples. However, it should be understood that the following examples are only for illustrative purposes and are not intended to limit the scope of the invention.

<Example 1>

Cloning of *Hansemula polymorpha* yapsin1 gene *HpYPS1* and analysis of the sequence

In order to find a gene encoding yapsin1 in *H. polymorpha*, a pair of synthetic oligonucleotides (5`-GAAGTGCAGCAGCAGCTCCTGAACC-3'; SEQ ID NO: 3, 5`-

GGCTGATGACGGCTCGGTCACGATGG-3'; SEQ ID NO: 4) were prepared, on the basis of information on *H. polymorpha* random sequenced tags described in Blandin et al., (FEBS Lett. 487, 76, (2000). By PCR (Polymerase Chain Reaction) using the prepared oligonucleotides as primers, a 0.88 kb DNA segment was amplified from *H. polymorpha* DL-1L derived chromosome. Then, Southern blotting was conducted using the amplified DNA segment as a probe. Based on the result of the Southern blot, 3.5 kb HindIII DNA segments was extracted from *H. polymorpha* chromosomal DNA to prepare a genome library, which was transformed into *E. coli.* A DNA segment reactive with the DNA probe was isolated by colony PCR and subjected to DNA sequencing to identify a DNA segment comprising ORF (Open Reading Frame) with a size of1728 bp showing a high similarity with the *Saccharomyces S. cerevisiae YPS1* gene (Fig. 1). The *H. polymorpha YPS1* gene (*HpYPS1*) product has a signal sequence of 1-17 amino acids at its N-end and a region of 556-575 amino acids, presumed as a domain which can be anchored on the membrane of Golgi apparatus, as reported on *S. cerevisiae* yapsin1.

Also, it has structural features, by which a glycosylphosphatidylinositol anchor can be attached thereto, as in *S.cerevisiae* yapsins (Fig. 1). *H. polymorpha* yapsin1 (SEQ ID NO: 2) shows a homology of 36% and a high similarity of 52% with *S. cerevisiae* yapsin1 and a homology of 30% or more with other yapsin proteases (Fig. 2, Table 1).

Table 1

Homology and similarity of *HpYPS1* and *ScYPS* genes

| | *ScYPS1* | *ScYPS2* | *ScYPS3* | *ScYPS6* | *ScYPS7* |
|---|---|---|---|---|---|
| *HpYPS1* | 36%(52%) | 31%(49%) | 30%(44%) | 26%(44%) | 29%(34%) |

* Parenthesized number is similarity

<Example 2>

Functional analysis of the *Hansenula polymorpha HpYPS1* gene

In case of the traditional yeast *Saccharomyces cerevisiae*, it has been reported that *YPS1/YPS2* double deficient mutant strain derived from *S. cerevisiae* W303 strain as a parental strain shows a such high temperature sensitivity that it cannot grow at 37℃ (Komano and Fuller, Proc. Natl. Acad. Sci. USA 7,92,10752-10756 (1995)). Also, SLH18(*yps1*△/*yps2*△/*yps3*△) strain (Korean Patent Registration No. 10-0386836), which is a yapsin multiple deficient mutant strain having three genes of *YPS1/YPS2/YPS3* destroyed, prepared using *Saccharomyces cerevisiae* L3262 as a parental strain by the present inventors, was found to show temperature sensitivity at a high temperature. As a method for functional analysis of the *Hansenula polymorpha* yapsin1 gene *HpYPS1*, a functional complementation experiment was conducted by transforming the *HpYPS1* gene into the above-described *Saccharomyces cerevisiae* multiple yapsin mutant strain according to the lithium chloride-DMSO method (Hill et al., Nucleic Acid Res., 19, 5791 (1991)) to examine whether the expression of *Hansenula polymorpha* yapsin1 gene could restore temperature sensitivities of the *Saccharomyces cerevisiae* multiple yapsin mutant

strains. In order to express the *HpYPS1* gene in *Saccharomyces cerevisiae*, 3.2 kb DNA segment including the *HpYPS1* gene was inserted into the *Saccharomyces cerevisiae* multicopy vector YEp352 (Hill et al., Yeast 2, 163, (1986)) to generate YEp-*HpYPS1*.

As shown in Fig. 3, the *Saccharomyces cerevisiae* yapsin multiple deficient mutant strains transformed with the YEp352-*HpYPS1* vector containing the *HpYPS1* gene showed the same growth with the wild type strain at 37℃, while the yapsin multiple deficient mutant strain transformed with only YEp352 vector, used as control, did not grow at 37℃. From these results, it was proved that the temperature sensitivity due to the multiple deletion of *Saccharomyces cerevisiae* yapsin genes can be overcome by the expression of the *Hansenula polymorpha* yapsin 1 gene. Thus, there was provided a basis supporting that the *H. polymorpha* gene *HpYPS1* cloned according to the present invention is a functional homologue of yapsin gene of *S. cerevisiae*.

<Example 3>

Preparation of the *HpYPS1* gene-deficient strain and analysis of yapsin activity

Using the *HpYPS1* gene prepared above, a *H. polymorpha* yapsin deficient mutant strain *hpyps1*△ was synthesized and assayed for change in its protein decomposition activity by yapsin. In order to prepare a mutant strain having the deletion of the *HpYPS1* gene encoding *Hansenula polymorpha* yapsin 1, a fusion PCR (Oldenhurg et al., Nucleic Acid Res. 25, 451, (1997)) was conducted on *H. polymorpha* chromosomal DNA using primers described in Table 2. The resulting DNA segment was transformed

into *Hansenula polymorpha* DL1-LdU (leu2△ura3::lacz; Kang et al., In *Hansenula polymorpha*:Biology and Application (Ed. G. Gellissen), pp 124 (2002)) strain to induce *in vivo* homologous recombination, thereby attempting deletion of the gene. More specifically, a primary PCR was firstly conducted to obtain the N-ends, C-ends of *HpYPS1* gene and URA3 gene. Then, a secondary fusion PCR was conducted to fuse the N-end and C-end of *HpYPS1* gene with the N-end and C-end of URA3 gene, respectively. The two DNA segments thus obtained were introduced to yeast cells.

Using *URA3* as a selection marker, viable transformants were selected in a minimal medium lacking uracil. Construction of the *hpyps1*△ strain (*leu2*△*ura3::lacz yps1:: URA3*), which is a *H. polymorpha* mutant strain having the deleted *HpYPS1* gene, was confirmed by PCR, which generated the DNA segments with different sizes between the wild type strain and the mutant strain. (Fig. 4).

For the comparison of the yapsin activity of *H. polymorpha hpyps1*△ mutant strain prepared above with that of the wild type strain, each of the two strains was cultured for 10 hours in YPD medium (1% yeast extract, 2% peptone, 2% glucose) using human parathyroid hormone (hPTH) as a substrate. The resulting yeast culture supernatant was analyzed for its protein degradation activity. 20 ㎕ of the yeast culture supernatant which had been diluted 1/4 was mixed with hPTH (about 1.6 ㎍) acting as a substrate. The reaction was incubated for 2 hours, 4 hours and 6 hours at 37℃. Each product was loaded on SDS-polyacrylamide gel to analyze the degradation level of hPTH by the yapsin activity.

As shown in Fig. 5, in the culture fluid of the wild type, it was already hard to observe hPTH remaining in the culture supernatant after 2 hour cultivation, while in the culture supernatant of *hpyps1*△ mutant strain, a substantial amount of hPTH was observed, though the amount was somewhat reduced as the reaction time became longer.

This suggests that the culture supernatant of *hpyps1*△, the *Hansenula polymorpha HpYPS1* deficient mutant strain, had a significantly reduced hPTH degradation activity, as compared to the culture supernatant of the wild type strain. Such result shows the same result with the previous study performed on *S. cerevisiae YPS1* deficient mutant strain (Kang et al., Appl Microbiol Biotechnol. 50, 187, (1998)) and thus, supports the fact that *HpYPS1* is a gene encoding yapsin protease of *H. polymorpha* and the *HpYPS1* deficient mutant strain developed according to the present invention has the yapsin activity reduced.

Table 2

Primers used in fusion PCR for destruction of *HpYPS1* gene

| Primer | Sequence | SEQ ID NO: |
|---|---|---|
| *YPS*(NF) | 5'-GGACACGCAAGAGGTGTCTG- 3' | 5 |
| *YPS*(NR+rp) | 5'-AGCTCGCTACCCGGGGATCCGCAACTTTCATTGTGTCAAC- 3' | 6 |
| *YPS*(CF+rp) | 5'-GCACATCCCCCTTTCGCCAGCCTCTTCGGTGCGGTTGACC- 3' | 7 |
| *YPS*(CR) | 5'-GCTCGGCTCCAGGATTCAGG- 3' | 8 |
| *URA3* N-S | 5'-GGATCCCCGGGTACCGAGCT- 3' | 9 |

| | | |
|---|---|---|
| *URA3* N-A | 5'-CACCGGTAGCTAATGATCCC- 3' | 10 |
| *URA3* C-S | 5'-CGAACATCCAAGTGGGCCGA- 3' | 11 |
| *URA3* C-A | 5'-CTGGCGAAAGGGGGATGTGC- 3' | 12 |

<Example 4>

Construction of the recombinant *H. polymorpha* strain expressing hPTH and analysis of the hPTH expression

In order to practically express and secret human parathyroid hormone (hPTH) in *H. polymorpha*, pMOXhPTH, a human parathyroid hormone expression vector for *H. polymorpha* was prepared according to a method shown in Fig. 6. Thus, the 0.53 kb EcoR I /Sal I segment containing the hPTH cDNA fused with a MFα signal sequence was prepared from pG10-hPTH , a hPTH expression vector for *S. cerevisiae* (Chung et al., Biotechnol Bioeng. 57, 245, (1998)) and the 7.8 kb Xba I /EcoR I segment having the albumin cDNA removed was prepared from pYHSA12, a human serum albumin expression vector for *H. polymorpha* (Kang et al., Biotechnol Bioeng. 76, 175, (2001)).

The two DNA segments were joined to generate pMOXhPTH. The resulting vector pMOXhPTH was introduced into the *H. polymorpha hpyps1*△ mutant strain and the wild type strain, and viable transformants were selected in minimal medium lacking leucine. At every 24 hours after inoculation, $Leu^+$ transformants were transferred to liquid minimal selective medium lacking leucine 5 times to stabilize the $Leu^+$ transformants (Sohn et al., Appl Microbiol Biotechnol. 51, 800, (1999)). The culture broth of the $Leu^+$ transformants obtained as described above was plated onto minimal

media containing G418 at various concentrations and cultured at 37℃. DNA was isolated from each of the resulting colonies and subjected to Southern blot analysis using the 1.5 kb *H. polymorpha LEU2* gene as a hybridization prove according to the method described by Sambrook et al. (Molecular cloning Cold Spring Harbor Laboratory Press, 1989) to confirm insertion into the chromosome of the expression vector (Fig. 7A). The *H. polymorpha LEU2* gene used as a probe was prepared by labeling with digoxigenin using the non-radioactive DNA labeling and detection kit. As having been expected, in the transformants selected on the medium containing a high G418 concentration, multiple integration of the vector pMOXhPTH was observed. Upon comparison of intensity between the *LEU2* gene signal on the chromosome and the *LEU2* gene signal on the inserted vector, it was presumed that about 5 to 6 copies at most had been inserted.

Some tranformants (a, b, c, d, e and f) of the recombinant yeast strains which had been confirmed to have the hPTH expression vector integrated into the chromosome were inoculated into YPM medium (1% yeast extract, 2% peptone, 2% methanol) and cultured at 37℃. Yeast culture supernatants obtained at 12 hours and 24 hours after initiation of the cultivation were treated with TCA (trichloroacetic acid). The proteins secreted out of the cells were concentrated to 1/20 of the initial volume and electrophoresed on 15% SDS-polyacrylamide gel, followed by staining with Coomassie Brilliant Blue R-250 (Fig. 7B).

When the hPTH secreted in *polymorpha* wild type was compared with that secreted in the *hpyps1*△ mutant strain, only a trace of decomposition product (d1) was

observed in the culture supernatant obtained from the wild type and hPTH of a whole size (i) was hardly observed, since most of hPTH was decomposed. On the other hand, secretion of hPTH of a whole size was clearly observed in the culture supernatant obtained from the *hpyps1*△ mutant strain at 12 hours after cultivation. From these results, it was shown that the *hpyps1*△ mutant strain having the yapsin1 gene destroyed is a more useful strain as a host for secretory expression of recombinant hPTH, as compared to the wild type, since the decomposition of the recombinant parathyroid hormone is considerably inhibited by the reduction of yapsin activity, although hPTH of a whole size was reduced in the *hpyps1*△ mutant strain while the band (d2), it is presumed, of hPTH having the C-end decomposed was observed, as the cultivation time became longer.

<Example 5>

Analysis of expression and secretion of recombinant human serum albumin in *H. polymorpha HpYPS1* gene deficient mutant strain

The human serum albumin (HSA) expressed in *S. cerevisiae* is secreted as an intact form of 67 kDa, however some recombinant HSA in a decomposed form of 45 kDa also have been observed. It was reported that degradation of HSA secreted to yeast cell culture supernatant, particularly production of a decomposition product with a size of 45 kDa was reduced when a *S. cerevisiae* strain with the yapsin1 gene destroyed was used as a host (Kerry Williams et al., Yeast 14, 161, (1998)). In order to analyze the expression

and degradation aspects of recombinant HSA in the *H. polymorpha HpYPS1* gene deficient mutant strain developed according to the present invention, the *hpyps1*△ mutant strain was transformed with an expression vector pYHSA12 (Kang et al., Biotech Bioeng. 76, 175, (2001)), in which the *MOX* promoter and the HSA cDNA were joined to a vector for copy-number controlled gene integration using the *LEU2* gene and G418 resistance gene as selective markers. Recombinant *H. polymorpha* wild type strain and *hpyps1*△ mutant strain were cultured in YPGM medium (1% yeast extract, 2% peptone, 1% glycerol, 2% methanol). Each of the culture supernatants thus-obtained was loaded on SDS-polyacrylamide gel, followed by staining with silver nitrate, or transferred to a nitrocellulose membrane, followed by Western blot by HSA antibody to examine expression and degradation aspects of HSA. As shown in Fig. 8, it was observed that more HSA was secreted in the *hpyps1*△ mutant strain than the wild type strain, particularly after cultivation for 24 hours. Though the overall degradation of albumin was not significantly inhibited, it was shown that degradation product with a size of 45 kDa was apparently reduced in the *hpyps1*△ mutant strain, as compared to the wild type, as observed particularly in *S. cerevisiae*. Considering cultivation in a high concentration fermenter, production of 45 kDa degraded HSA product presents more serious problems, it is expected that an albumin production system using the *hpyps1*△ mutant strain as a host can increase production of albumin due to a remarkable reduction in degradation of albumin, as compared to an production system using the wild type strain as a host.

<Example 6>

Analysis of expression and secretion of recombinant albumin fusion TIMP2 (HSA-TIMP2) in *H. polymorpha HpYPS1* gene deficient mutant strain

As a part of a method to increase blood persistence of a protein therapeutic agent, researches to develop a technology to increase *in vivo* half-life of a medical protein by expressing a recombinant protein in a form fused to albumin, a blood protein which has a long stability, are in the spotlight (Smith et al., Bioconjugate Chem. 12, 750-756, (2001); Sheffield et al. Blood Coagul Fibrinol. 12, 433-443, (2001)). The present inventors also have developed a recombinant TIMP-2 having *in vivo* stability significantly increased by expressing TIMP-2, which attracts public attention as a next generation anti-angiogenesis agent and anti-tumor agent, as a recombinant protein fused to albumin in *S. cerevisiae* (Korea Patent Registration No. 10-2001-0027823, International Application No. PCT/KR03/00015). In order to analyze the expression and e decomposition aspect of the recombinant HSA-TIMP2 in a form fused to albumin in the *H. polymorpha HpYPS1* gene deficient mutant strain (*hpyps1*△) developed according to the present invention, YHSA13-T2, a HSA-TIMP2 expression vector for *H. polymorpha* was prepared according to the method shown in Fig. 9. That is, the DNA fragments encoding HSA and TIMP2 were prepared using the PCR primers described in Table 3. Fusion PCR was conducted using the prepared genes in a ratio of 1:1 to prepare the 2.4 kb *HSA-TIMP2* DNA segment comprising HSA (1.8 kb) and TIMP2 (0.586 kb), which are connected to each other. The resulting segment was cloned into pGEM T vector

(Promega, USA) to prepare pTHSA-TIMP2. The sequence of the HSA-TIMP2 DNA segment was confirmed by sequencing analysis. Then, pTHSA-TIMP2 was cut with *Xba*I and *Spe*I. The resulting *Xba*I/*Spe*I HSA-TIMP2 gene segment of 1.2 kb was joined to the XbaI-digested pYHSA12(+), a *H. polymorpha* multiple tandem introduction vector having albumin gene inserted (Kang et al., *Biotechnol. Bioeng.* 76, 175-185, (2001)), to prepare pYHSA13-T2 (Fig. 9).

Table 3

PCR primers used in preparation of HSA-TIMP2 fusion gene

| primer | Sequence | Note |
|---|---|---|
| HSA EcoR F | 5' gaattcatgaagtgggtaaccttt 3' (SEQ ID NO: 13) | HSA forward direction |
| Hs-R2 | 5' taagcctaaggcagcttgac 3' (SEQ ID NO: 14) | HSA reverse direction |
| H-T2-F | 5'caagctgccttaggcttatgcagctgctccccggtg 3' (SEQ ID NO: 15) | Timp2 forward direction, used with 18 bp at HSA C-end in fusion PCR |
| R-T2-Sp | 5' actagtgatttatgggtcctcgatg 3' (SEQ ID NO: 16) | Timp2 reverse direction |

The albumin fusion protein expression vector prepared above was introduced into *H. polymorpha* DL1-L (leu2) and the *hpyps1*$\Delta$ *(leu2 hpyps1::URA3)* mutant strain to prepare $Leu^+$ transformants. The collected transformants were passage-cultured five times on selective medium so that the expression vectors could be multiply integrated to the host chromosomal DNA. Then, the transformants were plated at a density of $1x\ 10^5$ to $10^6$ per plate in media containing antibiotic G418 at various concentrations to select

transformants having resistance to G418. In order to examine whether the transformants, which had been confirmed to have the expression vector pYHSA13-T2 integrated into the host chromosome by Western blotting, secrets and expresses the fusion protein with an expected size, the yeast culture supernatants obtained by culturing the transformants for 48 hours in YPM was analyzed by Western blotting using antibody to albumin. As shown in Fig. 10, it was observed that the fusion protein HSA-TIMP2 (88 kDa) having an increased size as compared to HSA (66.5 kDa) was secreted as expected. Interestingly, in case of the yapsin deficient mutant strain *hpyps1*$\Delta$, only 88 kDa HSA-TIMP2 with a whole size was observed without any decomposition product. However, in the wild type DL1 strain, a band being presumed as a decomposition product was observed just under the band of HSA-TIMP2 with a size of 88 kDa. Therefore, due to the reduction of decomposition products, the expressed amount of HSA-Timp2 fusion protein was about two times higher in the *hpyps1*$\Delta$ mutant strain than the wild type strain. This suggests that protein degradation by yapsin1 is inhibited in secretion of recombinant proteins expressed as an albumin-fusion form as well as secretion of the above-described recombinant albumin in the *hpyps1*$\Delta$ mutant strain

**Industrial Applicability**

The present invention can be usefully used in bioengineering industry to produce a recombinant protein using *H. polymorpha* since cleavage of the recombinant protein by

yapsin activity can be remarkably reduced by using *H. polymorpha* strain (*hpyps1*△) with the protease yapsin1 gene being deficient as a recombinant protein-producing host, thereby secreting and producing the recombinant protein in an intact configuration at a high yield.

| Applicant's or agent's | International application No. |
|---|---|
| File reference PCTA/KRIB/1 | PCT/KR03/01279 |

INDICATIONS RELATING TO DEPOSITED MICROORGANISM

OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

5

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description on page 14, line 1-2

**B. IDENTIFICATION OF DEPOSIT** Further deposits are on an additional sheet ☐

Name of depositary institution

Korean Collection for Type Cultures

Address of depositary institution (*including postal code and country*)

#52, Oun-dong, Yusong-ku, Taejon 305-806,
Republic of Korea

| Date of deposit | Accession Number |
|---|---|
| 18/06/2002 | KCTC 10281BP |

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet ☐

D.DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E.SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

| For receiving Office use only | For international Bureau use only |
|---|---|
| ☐ This sheet was received with the international application | ☐ This sheet was received by the International Bureau on: |
| Authorized officer | Authorized officer |

| Applicant's or agent's File reference PCTA/KRIB/1 | International application No. PCT/KR03/01279 |
|---|---|

INDICATIONS RELATING TO DEPOSITED MICROORGANISM

OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

5

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description on page 14, line 5-6

**B. IDENTIFICATION OF DEPOSIT** Further deposits are on an additional sheet ☐

Name of depositary institution

Korean Collection for Type Cultures

Address of depositary institution(*including postal code and country*)

#52, Oun-dong, Yusong-ku, Taejon 305-806,
Republic of Korea

| Date of deposit | Accession Number |
|---|---|
| 18/06/2002 | KCTC 10282BP |

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet ☐

D.DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E.SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

| For receiving Office use only | For international Bureau use only |
|---|---|
| ☐ This sheet was received with the international application | ☐ This sheet was received by the International Bureau on: |
| Authorized officer | Authorized officer |

| Applicant's or agent's | International application No. |
|---|---|
| File reference PCTA/KRIB/1 | PCT/KR03/01279 |

INDICATIONS RELATING TO DEPOSITED MICROORGANISM

OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

5

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description on page 14, line 8-9

B. IDENTIFICATION OF DEPOSIT Further deposits are on an additional sheet ☐

Name of depositary institution

Korean Collection for Type Cultures

Address of depositary institution (*including postal code and country*)

#52, Oun-dong, Yusong-ku, Taejon 305-806,
Republic of Korea

| Date of deposit | Accession Number |
|---|---|
| 18/06/2002 | KCTC 10283BP |

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet ☐

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications e.g., "Accession Number of Deposit"*)

| For receiving Office use only | For international Bureau use only |
|---|---|
| ☐ This sheet was received with the international application | ☐ This sheet was received by the International Bureau on: |
| Authorized officer | Authorized officer |

| Applicant's or agent's File reference PCTA/KRIB/1 | International application No. PCT/KR03/01279 |
|---|---|

INDICATIONS RELATING TO DEPOSITED MICROORGANISM

OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

5

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description on page 13, line 18-20

**B. IDENTIFICATION OF DEPOSIT** Further deposits are on an additional sheet ☐

Name of depositary institution

Korean Collection for Type Cultures

Address of depositary institution (*including postal code and country*)

#52, Oun-dong, Yusong-ku, Taejon 305-806,
Republic of Korea

| Date of deposit | Accession Number |
|---|---|
| 18/06/2002 | KCTC 10285BP |

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet ☐

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications e.g., "Accession Number of Deposit"*)

| For receiving Office use only | For international Bureau use only |
|---|---|
| ☐ This sheet was received with the international application | ☐ This sheet was received by the International Bureau on: |
| Authorized officer | Authorized officer |

| Applicant's or agent's File reference PCTA/KRIB/1 | International application No. PCT/KR03/01279 |
|---|---|

INDICATIONS RELATING TO DEPOSITED MICROORGANISM

OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description on page 14, line 11-12

**B. IDENTIFICATION OF DEPOSIT** Further deposits are on an additional sheet ☐

Name of depositary institution

Korean Collection for Type Cultures

Address of depositary institution (*including postal code and country*)

#52, Oun-dong, Yusong-ku, Taejon 305-806,
Republic of Korea

| Date of deposit | Accession Number |
|---|---|
| 23/06/2002 | KCTC 10485BP |

C. ADDITIONAL INDICATIONS (*leave blank if not applicable*) This information is continued on an additional sheet ☐

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (*if the indications are not for all designated States*)

E. SEPARATE FURNISHING OF INDICATIONS (*leave blank if not applicable*)

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications e.g., "Accession Number of Deposit"*)

| For receiving Office use only | For international Bureau use only |
|---|---|
| ☐ This sheet was received with the international application | ☐ This sheet was received by the International Bureau on: |
| Authorized officer | Authorized officer |

**What Is Claimed Is:**

1. A nucleic acid molecule comprising the sequence shown in Fig. 1.

2. The nucleic acid molecule according to claim 1, which is the *HpYPS1* gene encoding *Hansenula polymorpha* yapsin1 (Accession No. KCTC 10285BP).

3. A polypeptide comprising the amino acid sequence shown in Fig.1.

4. The polypeptide according to claim 3, which is *Hansenula polymorpha* yapsin1 which can cleave a protein comprising a basic or dibasic amino acid residue.

5. The polypeptide according to claim 3, which is a secretion signal peptide of HpYPS1 polypeptide used as a secretion signal of a foreign protein.

6. A *Hansenula polymorpha* mutant strain having yapsin activity reduced by mutation of *HpYPS1* gene encoding *Hansenula polymorpha* yapsin1.

7. The *Hansenula polymorpha* mutant strain according to claim 6 deposited under Accession No. KCTC 10281BP.

8. A recombinant *Hansenula polymorpha* strain expressing a foreign protein which is prepared by introducing a gene encoding the foreign protein to the *Hansenula polymorpha* mutant strain according to claim 6.

9. The recombinant *Hansenula polymorpha* strain according to claim 8, which is *hpyps1* Δ-pMOXhPTH (KCTC 10282BP).

10. The recombinant *Hansenula polymorpha* strain according to claim 8, which is *hpyps1* Δ-pYHSA12 (KCTC 10283BP).

11. The recombinant *Hansenula polymorpha* strain according to claim 8, which is *hpyps1* Δ-pYHSA13-TIMP2 (KCTC 10485BP).

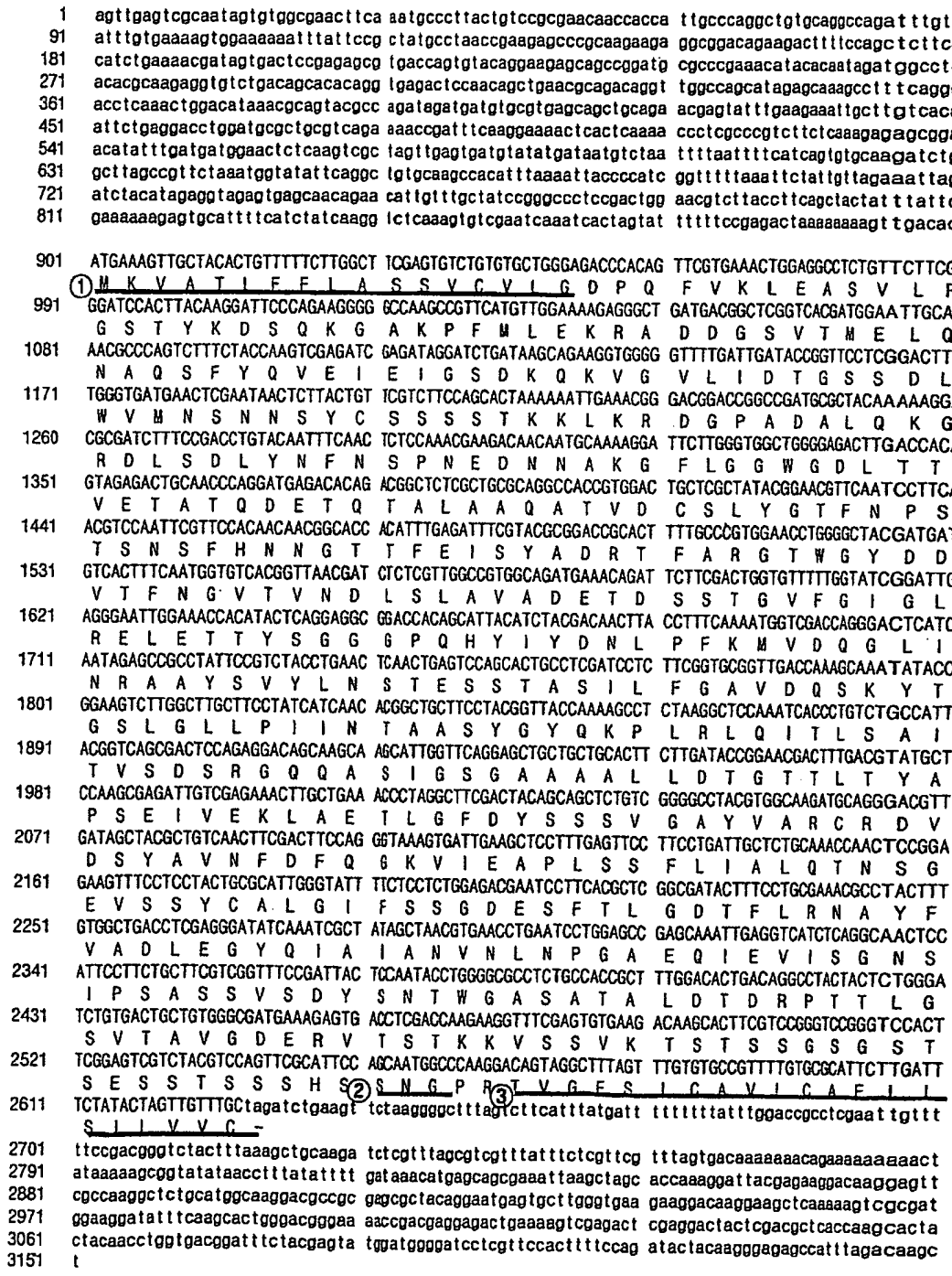
12. A process for preparing and isolating a foreign protein comprising expressing the foreign protein using the *Hansenula polymorpha* yapsin1 deficient strain according to any one of claims 8 to 11 as a host.

13. The process according to claim 12, in which the foreign protein is a recombinant protein containing a basic or dibasic amino acid residue which can be cleaved by yapsin1.

14. The process according to claim 13, in which the protein containing a basic or dibasic

amino acid residue comprises human parathyroid hormone, human serum albumin and albumin fusion protein.

# FIG. 1


1 agttgagtcgcaatagtgtggcgaacttca aatgcccttactgtccgcgaacaaccacca ttgcccaggctgtgcaggccagatttgtta
91 atttgtgaaaagtggaaaaaatttattccg ctatgcctaaccgaagagcccgcaagaaga ggcggacagaagacttttccagctcttcgg
181 catctgaaaacgatagtgactccgagagcg tgaccagtgtacaggaagagcagccggatg cgcccgaaacatacacaatagatggcctgg
271 acacgcaagaggtgtctgacagcacacagg tgagactccaacagctgaacgcagacaggt tggccagcatagagcaaagccttttcaggca
361 acctcaaactggacataaacgcagtacgcc agatagatgatgtgcgtgagcagctgcaga acgagtatttgaagaaattgcttgtcacat
451 attctgaggacctggatgcgctgcgtcaga aaaccgatttcaaggaaaactcactcaaaa ccctcgcccgtcttctcaaagagagcggaa
541 acatatttgatgatggaactctcaagtcgc tagttgagtgatgtatatgataatgtctaa ttttaattttcatcagtgtgcaagatctgg
631 gcttagccgttctaaatggtatattcaggc tgtgcaagccacatttaaaattaccccatc ggtttttaaattctattgttagaaattagg
721 atctacatagaggtagagtgagcaacagaa cattgtttgctatccgggccctccgactgg aacgtcttaccttcagctactatttattca
811 gaaaaaagagtgcattttcatctatcaagg tctcaaagtgtcgaatcaaatcactagtat tttttccgagactaaaaaaaagttgacaca

901 ATGAAAGTTGCTACACTGTTTTTCTTGGCT TCGAGTGTCTGTGTGCTGGGAGACCCACAG TTCGTGAAACTGGAGGCCTCTGTTCTTCGG
① M K V A T I F F I A S S V C V I G D P Q F V K L E A S V L R
991 GGATCCACTTACAAGGATTCCCAGAAGGGG GCCAAGCCGTTCATGTTGGAAAAGAGGGCT GATGACGGCTCGGTCACGATGGAATTGCAG
G S T Y K D S Q K G A K P F M L E K R A D D G S V T M E L Q
1081 AACGCCCAGTCTTTCTACCAAGTCGAGATC GAGATAGGATCTGATAAGCAGAAGGTGGGG GTTTTGATTGATACCGGTTCCTCGGACTTG
N A Q S F Y Q V E I E I G S D K Q K V G V L I D T G S S D L
1171 TGGGTGATGAACTCGAATAACTCTTACTGT TCGTCTTCCAGCACTAAAAAATTGAAACGG GACGGACCGGCCGATGCGCTACAAAAAGGA
W V M N S N N S Y C S S S S T K K L K R D G P A D A L Q K G
1260 CGCGATCTTTCCGACCTGTACAATTTCAAC TCTCCAAACGAAGACAACAATGCAAAAGGA TTCTTGGGTGGCTGGGGAGACTTGACCACA
R D L S D L Y N F N S P N E D N N A K G F L G G W G D L T T
1351 GTAGAGACTGCAACCCAGGATGAGACACAG ACGGCTCTCGCTGCGCAGGCCACCGTGGAC TGCTCGCTATACGGAACGTTCAATCCTTCA
V E T A T Q D E T Q T A L A A Q A T V D C S L Y G T F N P S
1441 ACGTCCAATTCGTTCCACAACAACGGCACC ACATTTGAGATTTCGTACGCGGACCGCACT TTTGCCCGTGGAACCTGGGGCTACGATGAT
T S N S F H N N G T T F E I S Y A D R T F A R G T W G Y D D
1531 GTCACTTTCAATGGTGTCACGGTTAACGAT CTCTCGTTGGCCGTGGCAGATGAAACAGAT TCTTCGACTGGTGTTTTTGGTATCGGATTG
V T F N G V T V N D L S L A V A D E T D S S T G V F G I G L
1621 AGGGAATTGGAAACCACATACTCAGGAGGC GGACCACAGCATTACATCTACGACAACTTA CCTTTCAAAATGGTCGACCAGGGACTCATC
R E L E T T Y S G G G P Q H Y I Y D N L P F K M V D Q G L I
1711 AATAGAGCCGCCTATTCCGTCTACCTGAAC TCAACTGAGTCCAGCACTGCCTCGATCCTC TTCGGTGCGGTTGACCAAAGCAAATATACC
N R A A Y S V Y L N S T E S S T A S I L F G A V D Q S K Y T
1801 GGAAGTCTTGGCTTGCTTCCTATCATCAAC ACGGCTGCTTCCTACGGTTACCAAAAGCCT CTAAGGCTCCAAATCACCCTGTCTGCCATT
G S L G L L P I I N T A A S Y G Y Q K P L R L Q I T L S A I
1891 ACGGTCAGCGACTCCAGAGGACAGCAAGCA AGCATTGGTTCAGGAGCTGCTGCTGCACTT CTTGATACCGGAACGACTTTGACGTATGCT
T V S D S R G Q Q A S I G S G A A A A L L D T G T T L T Y A
1981 CCAAGCGAGATTGTCGAGAAACTTGCTGAA ACCCTAGGCTTCGACTACAGCAGCTCTGTC GGGGCCTACGTGGCAAGATGCAGGGACGTT
P S E I V E K L A E T L G F D Y S S S V G A Y V A R C R D V
2071 GATAGCTACGCTGTCAACTTCGACTTCCAG GGTAAAGTGATTGAAGCTCCTTTGAGTTCC TTCCTGATTGCTCTGCAAACCAACTCCGGA
D S Y A V N F D F Q G K V I E A P L S S F L I A L Q T N S G
2161 GAAGTTTCCTCCTACTGCGCATTGGGTATT TTCTCCTCTGGAGACGAATCCTTCACGCTC GGCGATACTTTCCTGCGAAACGCCTACTTT
E V S S Y C A L G I F S S G D E S F T L G D T F L R N A Y F
2251 GTGGCTGACCTCGAGGGATATCAAATCGCT ATAGCTAACGTGAACCTGAATCCTGGAGCC GAGCAAATTGAGGTCATCTCAGGCAACTCC
V A D L E G Y Q I A I A N V N L N P G A E Q I E V I S G N S
2341 ATTCCTTCTGCTTCGTCGGTTTCCGATTAC TCCAATACCTGGGGCGCCTCTGCCACCGCT TTGGACACTGACAGGCCTACTACTCTGGGA
I P S A S S V S D Y S N T W G A S A T A L D T D R P T T L G
2431 TCTGTGACTGCTGTGGGCGATGAAAGAGTG ACCTCGACCAAGAAGGTTTCGAGTGTGAAG ACAAGCACTTCGTCCGGGTCCGGGTCCACT
S V T A V G D E R V T S T K K V S S V K T S T S S G S G S T
2521 TCGGAGTCGTCTACGTCCAGTTCGCATTCC AGCAATGGCCCAAGGACAGTAGGCTTTAGT TTGTGTGCCGTTTTGTGCGCATTCTTGATT
S E S S T S S S H S ② S N G P R ③ T V G F S I C A V I C A F I I
2611 TCTATACTAGTTGTTTGCtagatctgaagt tctaaggggctttagtcttcatttatgatt ttttttatttggaccgcctcgaattgttt
S I I V V C -
2701 ttccgacgggtctactttaaagctgcaaga tctcgtttagcgtcgtttatttctcgttcg tttagtgacaaaaaaacagaaaaaaaaact
2791 ataaaaagcggtatataaccttatatttt gataaacatgagcagcgaaattaagctagc accaaaggattacgagaaggacaaggagtt
2881 cgccaaggctctgcatggcaaggacgccgc gagcgctacaggaatgagtgcttgggtgaa gaaggacaaggaagctcaaaaagtcgcgat
2971 ggaaggatatttcaagcactgggacgggaa aaccgacgaggagactgaaaagtcgagact cgaggactactcgacgctcaccaagcacta
3061 ctacaacctggtgacggatttctacgagta tggatggggatcctcgttccacttttccag atactacaagggagagccatttagacaagc
3151 t

FIG. 2

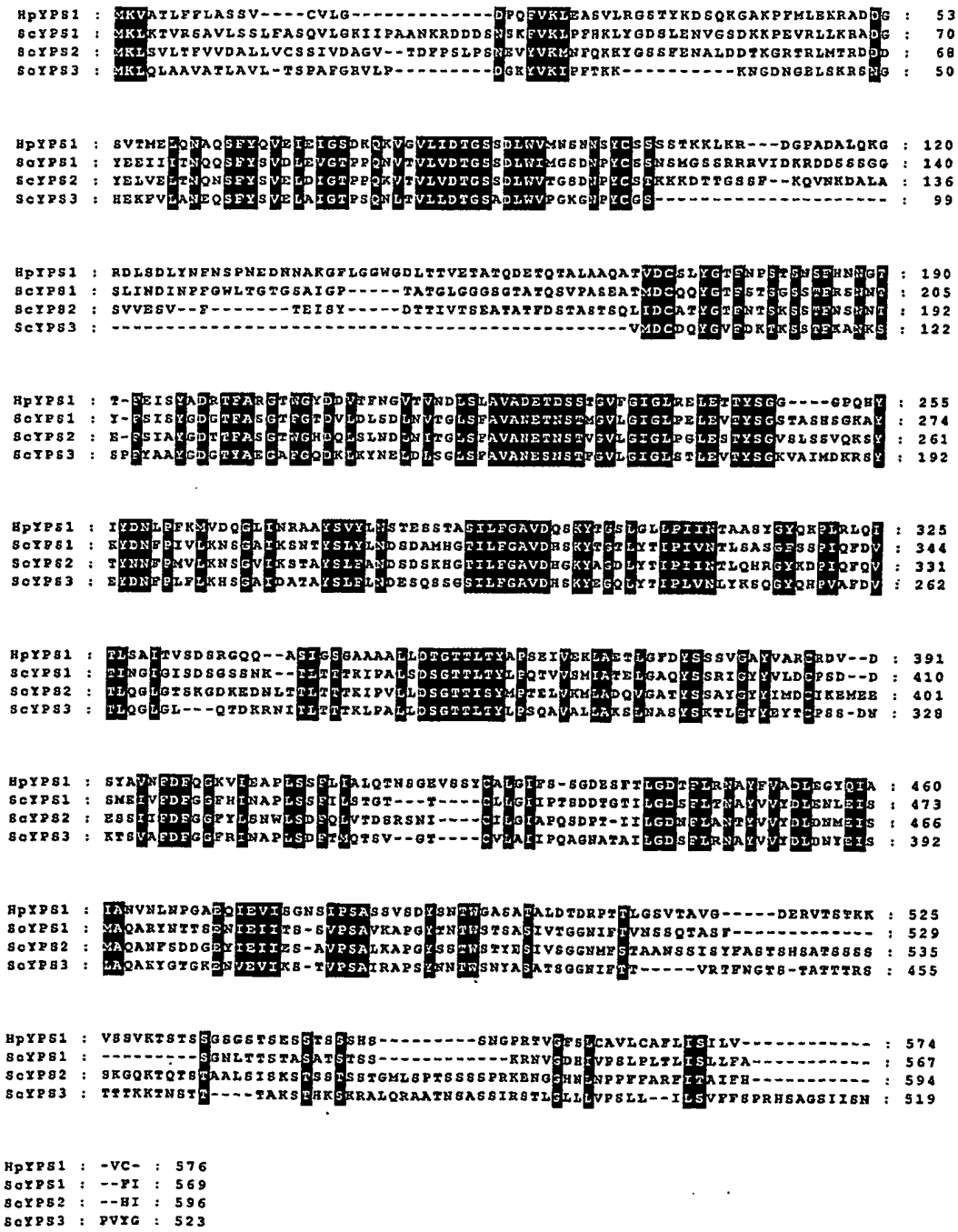

HpYPS1 : MKVATLFFLASSV----CVLG----------DPQFVKLEASVLRGSTYKDSQKGAKPFMLEKRADQG : 53
ScYPS1 : MKLKTVRSAVLSSLFASQVLGKIIPAANKRDDDSNSKFVKLPFHKLYGDSLENVGSDKKPEVRLLKRADG : 70
ScYPS2 : MKLSVLTFVVDALLVCSSIVDAGV--TDFPSLPSNEVFVKRMNFQKEYGSSFENALDDTKGRTRLMTRDQD : 68
ScYPS3 : MKLQLAAVATLAVL-TSPAFGRVLP---------DGKFVKIPFTKK----------KNGDNGELSKRSNG : 50
HpYPS1 : SVTMEFQNAQSFYQVEIEIGSDKQKVGVLIDTGSSDLWVMNSNNSYCSSSSTKKLKR---DGPADALQKG : 120
ScYPS1 : YEEIILTNQQSFYSVDLEVGTPPQNVTVLVDTGSSDLWIMGSDNPYCSSNSMGSSRRRVIDKRDDSSSGG : 140
ScYPS2 : YELVELTNQNSFYSVELDIGTPPQKVTVLVDTGSSDLWVTGSDNPYCSTKKKDTTGSSF--KQVNKDALA : 136
ScYPS3 : HEKFVLANEQSFYSVELAIGTPSQNLTVLLDTGSADLWVPGKGNPYCGS---------- : 99
HpYPS1 : RDLSDLYNFNSPNEDNNAKGFLGGWGDLTTVETATQDETQTALAAQAIVDCSLYGTSNPSTGNSFHNNGT : 190
ScYPS1 : SLINDINPFGWLTGTGSAIGP-----TATGLGGGSGTATQSVPASEATMDCQQYGTSSTSGSSTFRSNNT : 205
ScYPS2 : SVVESV--F--------TEISY-----DTTIVTSEATATFDSTASTSQLIDCATYGTFNTSKSSTFNSNNI : 192
ScYPS3 : ----------VMDCDQYGVFDKTKSSTFKANKS : 122
HpYPS1 : T-NEISNADRTFARGTNGYDDVTFNGVTVNDLSLAVADETDSSIGVFGIGLKELETTYSGG----GPQHY : 255
ScYPS1 : Y-GSISYGDGTFASGTFGTDVLDLSDLNVTGLSFAVANETNSTMGVLGIGLPELEVTYSGSTASHSGKAY : 274
ScYPS2 : E-GSIAYGDTTFASGTNGHDQLSLNDLNITGLSFAVANETNSTVGVLGIGLPGLESTYSGVSLSSVQKSY : 261
ScYPS3 : SPSYAAYGDGTYAEGAFGQDKLKYNELDLSGLSFAVANESNSTFGVLGIGLSTLEVTYSGKVAIMDKRSY : 192
HpYPS1 : IVDNLPFKLVDQGLINRAAYSVYLNSTESSTASILFGAVDQSKYTGSLGLLPIINTAASYSYQKPLRLQT : 325
ScYPS1 : EVDNFPIVLKNSGAIKSNTYSLYLNDSDAMHGTILFGAVDRSKYTGTLYTIPIVNTLSASGFSSPIQFDV : 344
ScYPS2 : TYNNFPMVLKNSGVIKSTAYSLFANDSDSKHGTILFGAVDEGKYAGDLYTIPIINTLQHRGYKDPIQFQV : 331
ScYPS3 : EVDNFPLFLKHSGAIDATAYSLFLNDESQSSGSILFGAVDHSKYEGQLYTIPLVNLYKSQGYQHEVAFDV : 262
HpYPS1 : TLSALTVSDSRGQQ--ASIGSGAAAALLDTGTTLTYAPSEIVSKLAETLGFDYSSSVGAYVARCRDV--D : 391
ScYPS1 : TINGIGISDSGSSNK--ILTTTKIPALSDSGTTLTYLPQTVVSMIATELGAQYSSRIGYYVLDCFSD--D : 410
ScYPS2 : TLQGLGTSKGDKEDNLTTLTTKIPVLLDSGTTISYMPTELVKMLADQVGATYSSAYGYYIMDCIKEMEE : 401
ScYPS3 : TLQGLGL---QTDKRNILTTTKLPALLDSGTTLTYLPSQAVALLAKSLNASYSKTLGYYEYTCPSS-DN : 328
HpYPS1 : SYAVNFDFQGKVISAPLSSGLIALQTNSGEVSSYCALGIFS-SGDESFTLGDTFLRNAYFVADLEGYQIA : 460
ScYPS1 : SMEIVFDFGGFHINAPLSSFILSTGT---T-----CLLGIIPTSDDTGTILGDSFLTNAYVVYDLENLEIS : 473
ScYPS2 : ESSIIFDFGGFPYLSNWLSDSQLVTDSRSNI-----CILGIAPQSDFT-IILGDNFLANTYVVYDLDNMEIS : 466
ScYPS3 : KTSVAFDFGGFRFNAPLSDFTMQTSV--GT-----CVLAIIPQAGNATAILGDSFLRNAYVVYDLDNYEIS : 392
HpYPS1 : LANVNLNPGAEQIEVISGNSIPSASSVSDKSNTWGASATALDTDRPTTLGSVTAVG-----DERVTSTKK : 525
ScYPS1 : MAQARYNTTSENIEIITS-GVPSAVKAPGYTNTWSTSASIVTGGNIFTVNSSQTASF---------- : 529
ScYPS2 : MAQANFSDDGEYIEIIES-AVPSALKAPGYSSTWSTYESIVSGGNMFSTAANSSISYFASTSHSATSSSS : 535
ScYPS3 : LAQAEYGTGKENVEVIKS-TVPSAIRAPSYTNTWSNYASATSGGNIFTT-----VRTFNGTS-TATTTRS : 455
HpYPS1 : VSSVKTSTSSGSGSTSESSTSPSHS----------SNGPRTVGFSLCAVLCAFLISILV---------- : 574
ScYPS1 : ----------SGNLTTSTASATTSS-----------KRNVGDHIVPSLPLTLISLLFA---------- : 567
ScYPS2 : SKGQKTQTSTAALSISKSTSSTSSTGMLSPTSSSSPRKENGSHNENPFFFARFITAIFH---------- : 594
ScYPS3 : TTTKKTNSTT----TAKSTHKGRRALQRAATNSASSIRSTLGLLLVPSLL--ILSVFFSPRHSAGSIISN : 519
HpYPS1 : -VC- : 576
ScYPS1 : --FI : 569
ScYPS2 : --HI : 596
ScYPS3 : PVYG : 523

FIG. 3

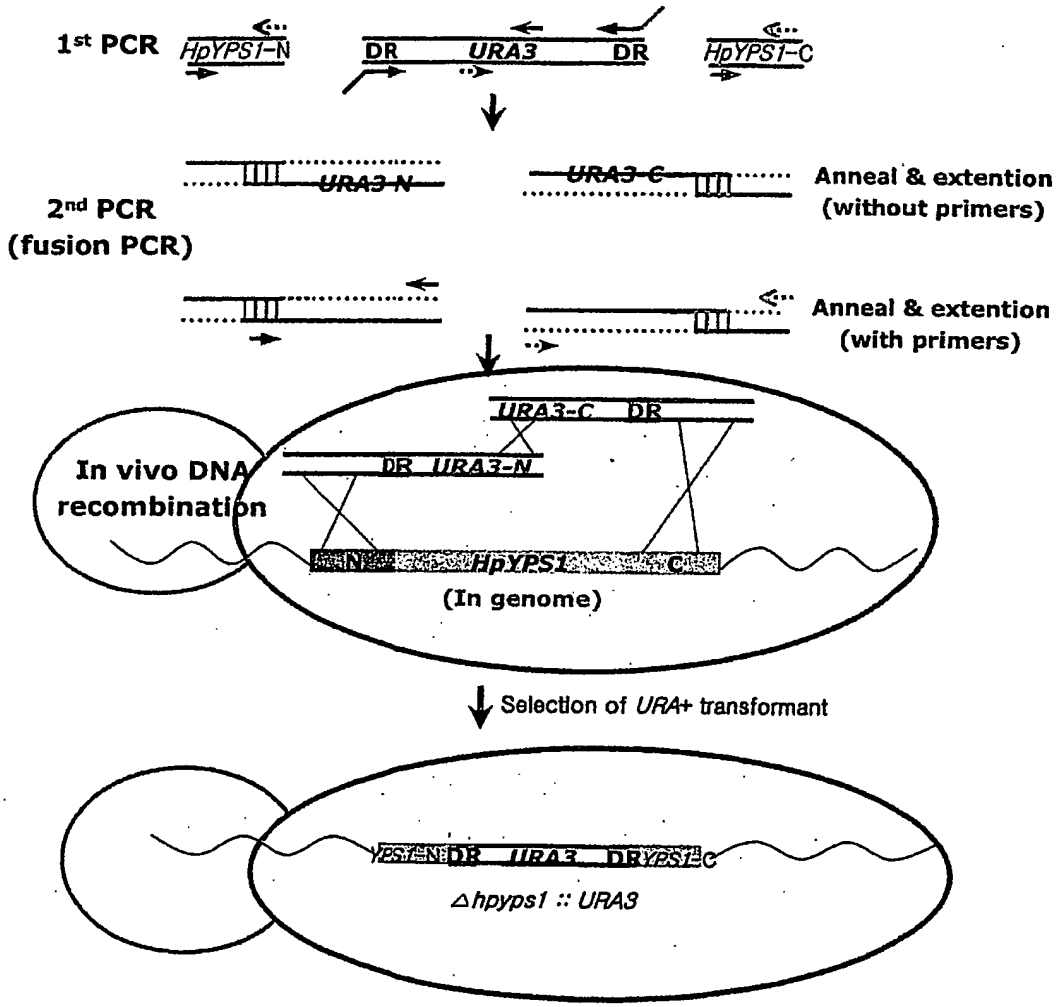

1st PCR
HpYPS1-N
DR
URA3
DR
HpYPS1-C
2nd PCR
(fusion PCR)
URA3-N
URA3-C
Anneal & extention
(without primers)
Anneal & extention
(with primers)
URA3-C DR
DR URA3-N
In vivo DNA
recombination
N
HpYPS1
C
(In genome)
Selection of URA+ transformant
YPS1-N DR URA3 DR YPS1-C
△hpyps1 :: URA3

FIG. 4

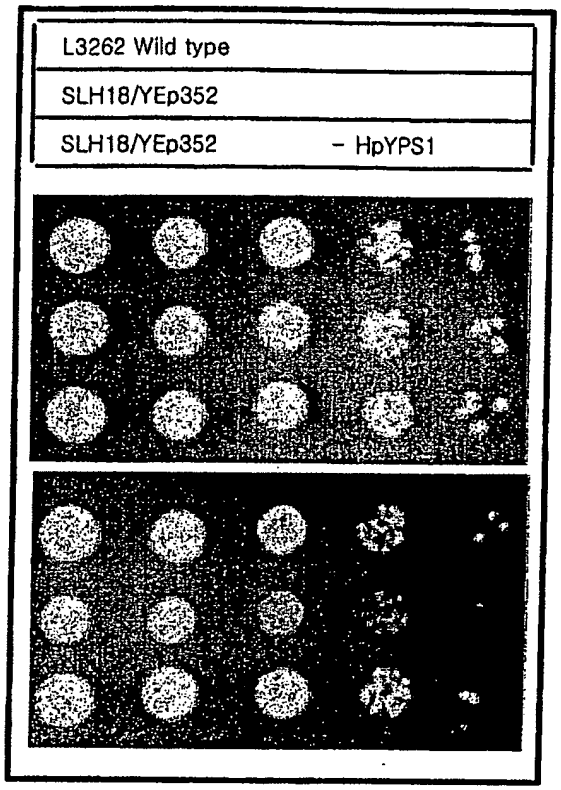
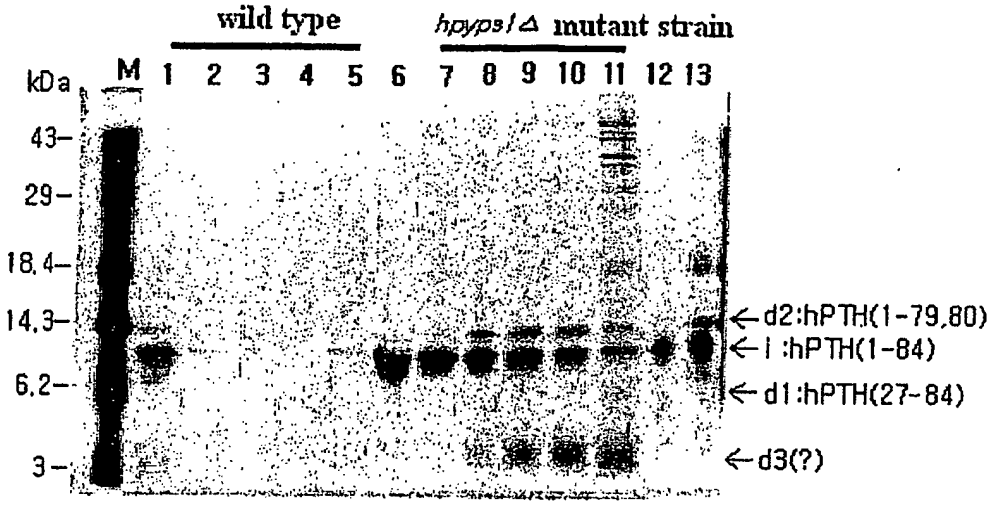


FIG. 5




Lane M: Molecular marker
1: wild type, o-h reaction
2: wild type, 2-h reaction
3: wild type, 4-h reaction
4: wild type, 6-h reaction
5: wild type, 24-h reaction
6: distilled water + hPTH, o-h reaction
7: mutant strain, o-h reaction
8: mutant strain, 2-h reaction
9: mutant strain, 4-h reaction
10: mutant strain, 6-h reaction
11: mutant strain, 24-h reaction
12: hPTH 100 ng
13: hPTH 200 ng

6/10
FIG. 6

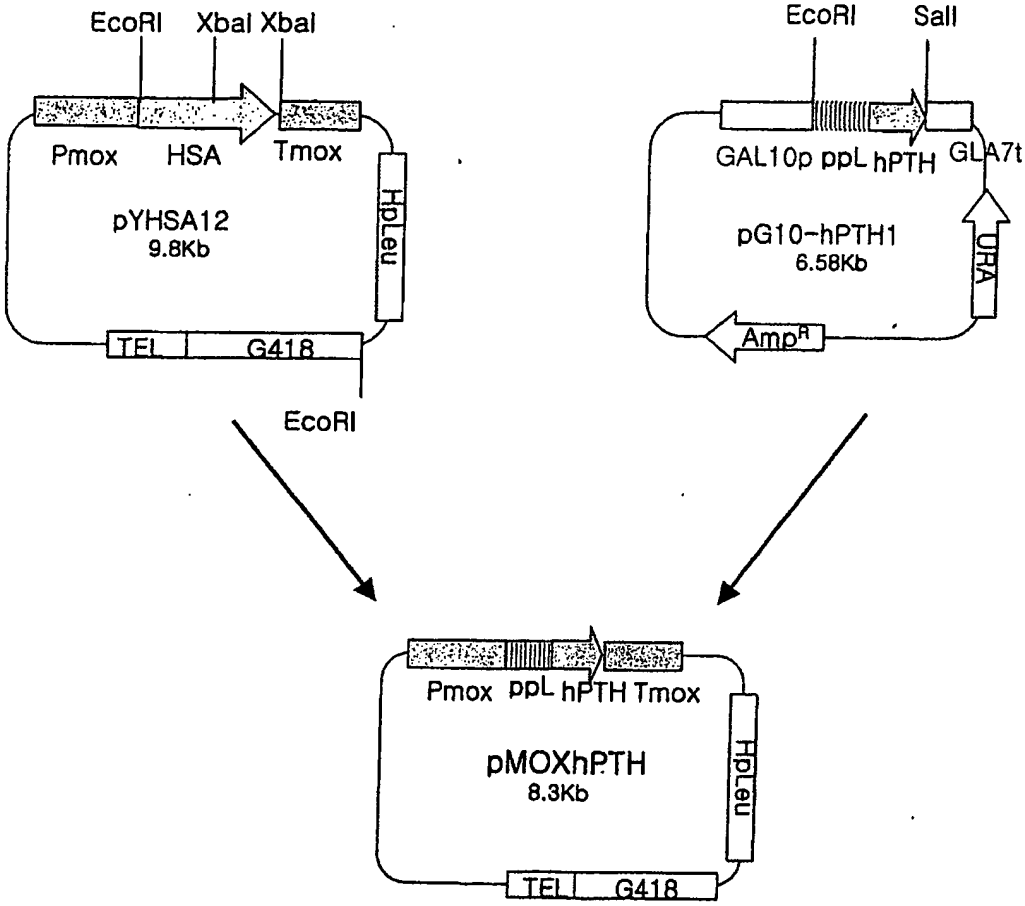
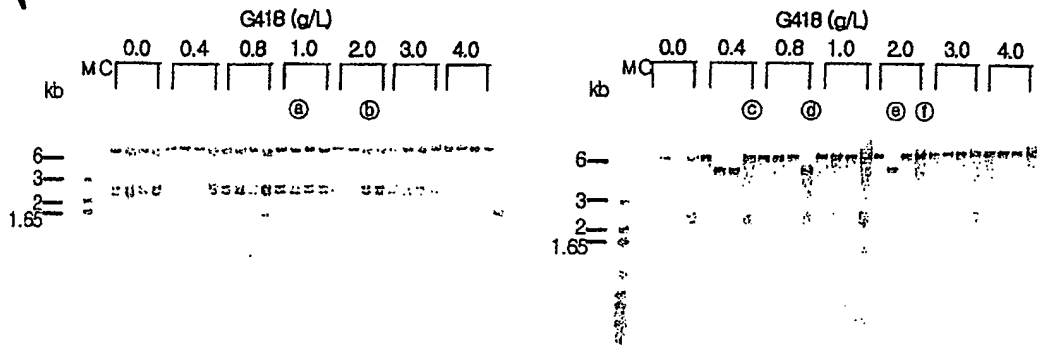

EcoRI
XbaI XbaI
Pmox
HSA
Tmox
pYHSA12
9.8Kb
HpLeu
TEL
G418
EcoRI
EcoRI
SalI
GAL10p ppL hPTH
GLA7t
pG10-hPTH1
6.58Kb
URA
Amp^R
Pmox ppL hPTH Tmox
pMOXhPTH
8.3Kb
HpLeu
TEL
G418

FIG. 7

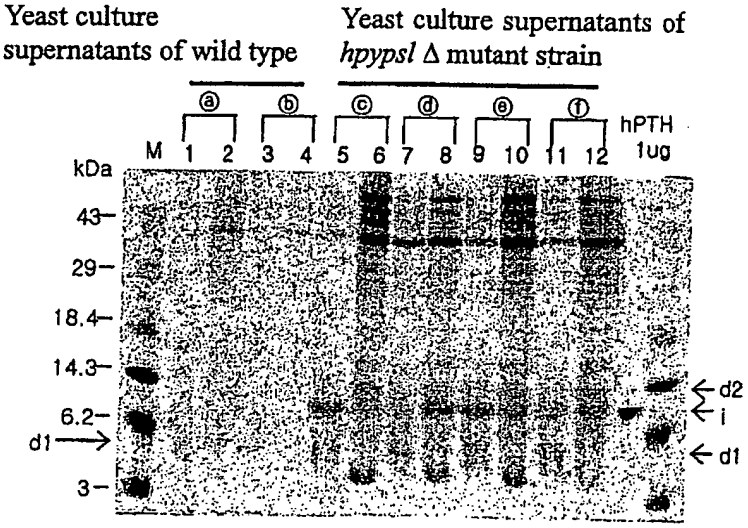
A




Transformants of wild type

Transformants of *hypysl* Δ mutant strain

B




Lane 1, 3, 5, 7, 9, 11 : 12hr after initiation of the cultivation
Lane 2, 4, 6, 8, 10, 12 : 24hr after initiation of the cultivation

FIG. 8

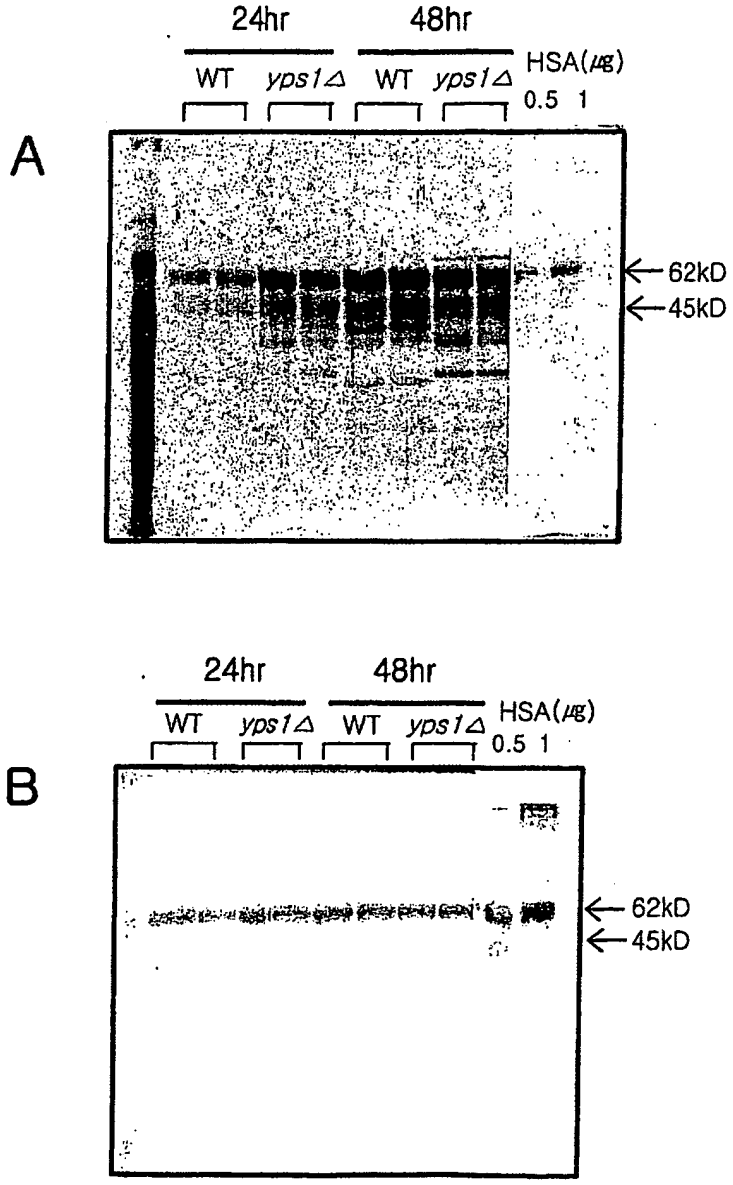

A
24hr
48hr
WT
yps1Δ
WT
yps1Δ
HSA(μg)
0.5 1
62kD
45kD
B
24hr
48hr
WT
yps1Δ
WT
yps1Δ
HSA(μg)
0.5 1
62kD
45kD

FIG. 9

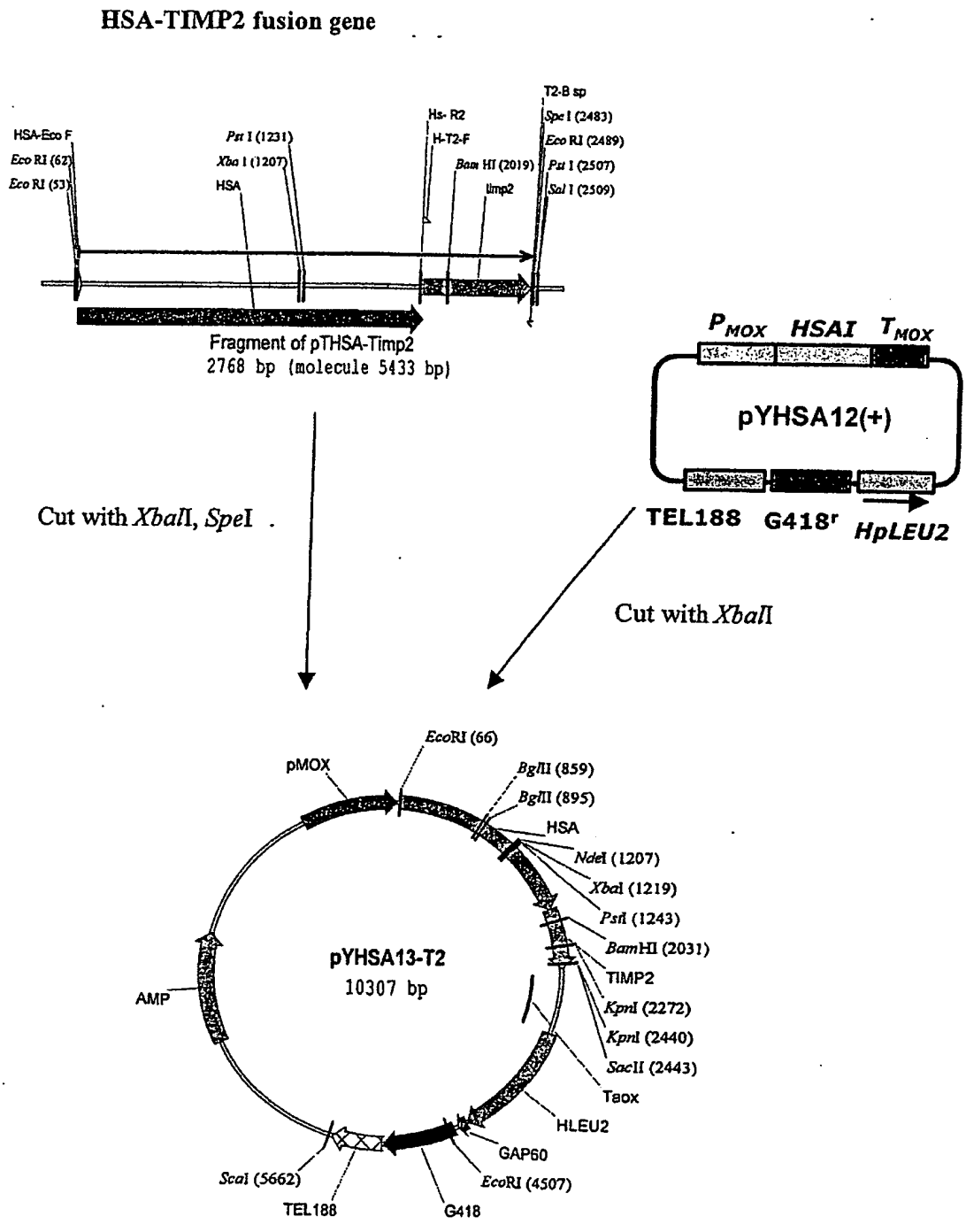
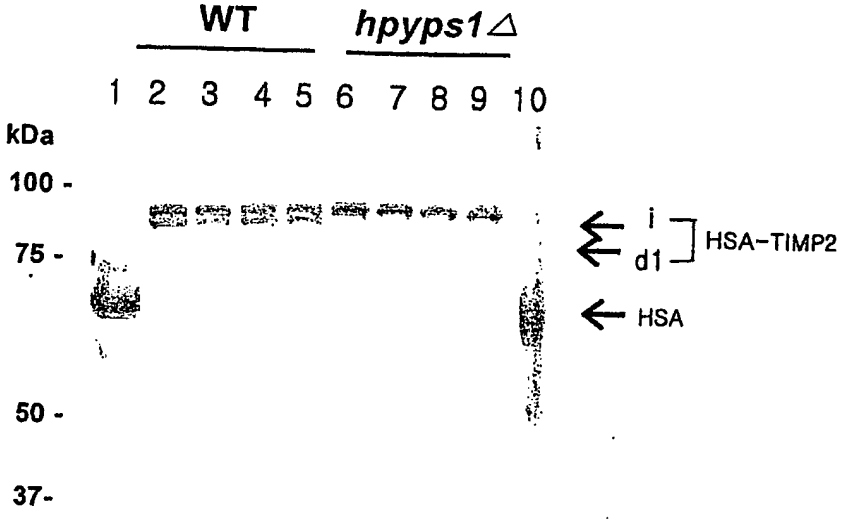

HSA-TIMP2 fusion gene
HSA-Eco F
Eco RI (62)
Eco RI (53)
Pst I (1231)
Xba I (1207)
HSA
Hs- R2
H-T2-F
Bam HI (2019)
timp2
T2-B sp
Spe I (2483)
Eco RI (2489)
Pst I (2507)
Sal I (2509)
Fragment of pTHSA-Timp2
2768 bp (molecule 5433 bp)
Cut with XbaII, SpeI
PMOX HSAI TMOX
pYHSA12(+)
TEL188 G418r HpLEU2
Cut with XbaII
pYHSA13-T2
10307 bp
pMOX
EcoRI (66)
BglII (859)
BglII (895)
HSA
NdeI (1207)
XbaI (1219)
PstI (1243)
BamHI (2031)
TIMP2
KpnI (2272)
KpnI (2440)
SacII (2443)
Taox
HLEU2
GAP60
EcoRI (4507)
G418
TEL188
ScaI (5662)
AMP

FIG. 10



# Sequence Listing

<110> Korea Research Institute of Bioscience and Biotechnology
LeadBio,Inc
Bio Holdings CO., LTD

<120> Hansenula polymorpha yapsin deficient mutant strain and process for the preparation of recombinant proteins using the same

<160> 16

<170> KopatentIn 1.71

<210> 1
<211> 3151
<212> DNA
<213> Hansenula polymorpha

<220>
<221> sig_peptide
<222> (901)..(903)
<223> initiation codon

<220>
<221> 5'UTR
<222> (1)..(900)

<220>
<221> 3'UTR
<222> (2622)..(3151)

<220>
<221> terminator
<222> (2629)..(2631)
<223> termination codon

<220>
<221> CDS
<222> (901)..(2628)
<223> coding sequence

# Sequence Listing

<400> 1

agttgagtcg caatagtgtg gcgaacttca aatgccctta ctgtccgcga acaaccacca 60

ttgcccaggc tgtgcaggcc agatttgtta atttgtgaaa agtggaaaaa atttattccg 120

ctatgcctaa ccgaagagcc cgcaagaaga ggcggacaga agacttttcc agctcttcgg 180

catctgaaaa cgatagtgac tccgagagcg tgaccagtgt acaggaagag cagccggatg 240

cgcccgaaac atacacaata gatggcctgg acacgcaaga ggtgtctgac agcacacagg 300

tgagactcca acagctgaac gcagacaggt tggccagcat agagcaaagc ctttcaggca 360

acctcaaact ggacataaac gcagtacgcc agatagatga tgtgcgtgag cagctgcaga 420

acgagtattt gaagaaattg cttgtcacat attctgagga cctggatgcg ctgcgtcaga 480

aaaccgattt caaggaaaac tcactcaaaa ccctcgcccg tcttctcaaa gagagcggaa 540

acatatttga tgatggaact ctcaagtcgc tagttgagtg atgtatatga taatgtctaa 600

ttttaatttt catcagtgtg caagatctgg gcttagccgt tctaaatggt atattcaggc 660

tgtgcaagcc acatttaaaa ttaccccatc ggtttttaaa ttctattgtt agaaattagg 720

atctacatag aggtagagtg agcaacagaa cattgtttgc tatccgggcc ctccgactgg 780

aacgtcttac cttcagctac tatttattca gaaaaaagag tgcattttca tctatcaagg 840

tctcaaagtg tcgaatcaaa tcactagtat tttttccgag actaaaaaaa agttgacaca 900

atg aaa gtt gct aca ctg ttt ttc ttg gct tcg agt gtc tgt gtg ctg 948
Met Lys Val Ala Thr Leu Phe Phe Leu Ala Ser Ser Val Cys Val Leu
1 5 10 15

gga gac cca cag ttc gtg aaa ctg gag gcc tct gtt ctt cgg gga tcc 996
Gly Asp Pro Gln Phe Val Lys Leu Glu Ala Ser Val Leu Arg Gly Ser
20 25 30

# Sequence Listing

```
act tac aag gat tcc cag aag ggg gcc aag ccg ttc atg ttg gaa aag        1044
Thr Tyr Lys Asp Ser Gln Lys Gly Ala Lys Pro Phe Met Leu Glu Lys
        35                  40                  45

agg gct gat gac ggc tcg gtc acg atg gaa ttg cag aac gcc cag tct        1092
Arg Ala Asp Asp Gly Ser Val Thr Met Glu Leu Gln Asn Ala Gln Ser
    50                  55                  60

ttc tac caa gtc gag atc gag ata gga tct gat aag cag aag gtg ggg        1140
Phe Tyr Gln Val Glu Ile Glu Ile Gly Ser Asp Lys Gln Lys Val Gly
65                  70                  75                  80

gtt ttg att gat acc ggt tcc tcg gac ttg tgg gtg atg aac tcg aat        1188
Val Leu Ile Asp Thr Gly Ser Ser Asp Leu Trp Val Met Asn Ser Asn
                85                  90                  95

aac tct tac tgt tcg tct tcc agc act aaa aaa ttg aaa cgg gac gga        1236
Asn Ser Tyr Cys Ser Ser Ser Ser Thr Lys Lys Leu Lys Arg Asp Gly
            100                 105                 110

ccg gcc gat gcg cta caa aaa gga cgc gat ctt tcc gac ctg tac aat        1284
Pro Ala Asp Ala Leu Gln Lys Gly Arg Asp Leu Ser Asp Leu Tyr Asn
        115                 120                 125

ttc aac tct cca aac gaa gac aac aat gca aaa gga ttc ttg ggt ggc        1332
Phe Asn Ser Pro Asn Glu Asp Asn Asn Ala Lys Gly Phe Leu Gly Gly
    130                 135                 140

tgg gga gac ttg acc aca gta gag act gca acc cag gat gag aca cag        1380
Trp Gly Asp Leu Thr Thr Val Glu Thr Ala Thr Gln Asp Glu Thr Gln
145                 150                 155                 160

acg gct ctc gct gcg cag gcc acc gtg gac tgc tcg cta tac gga acg        1428
Thr Ala Leu Ala Ala Gln Ala Thr Val Asp Cys Ser Leu Tyr Gly Thr
                165                 170                 175

ttc aat cct tca acg tcc aat tcg ttc cac aac aac ggc acc aca ttt        1476
Phe Asn Pro Ser Thr Ser Asn Ser Phe His Asn Asn Gly Thr Thr Phe
            180                 185                 190
```

# Sequence Listing

gag att tcg tac gcg gac cgc act ttt gcc cgt gga acc tgg ggc tac 1524
Glu Ile Ser Tyr Ala Asp Arg Thr Phe Ala Arg Gly Thr Trp Gly Tyr
195 200 205

gat gat gtc act ttc aat ggt gtc acg gtt aac gat ctc tcg ttg gcc 1572
Asp Asp Val Thr Phe Asn Gly Val Thr Val Asn Asp Leu Ser Leu Ala
210 215 220

gtg gca gat gaa aca gat tct tcg act ggt gtt ttt ggt atc gga ttg 1620
Val Ala Asp Glu Thr Asp Ser Ser Thr Gly Val Phe Gly Ile Gly Leu
225 230 235 240

agg gaa ttg gaa acc aca tac tca gga ggc gga cca cag cat tac atc 1668
Arg Glu Leu Glu Thr Thr Tyr Ser Gly Gly Gly Pro Gln His Tyr Ile
245 250 255

tac gac aac tta cct ttc aaa atg gtc gac cag gga ctc atc aat aga 1716
Tyr Asp Asn Leu Pro Phe Lys Met Val Asp Gln Gly Leu Ile Asn Arg
260 265 270

gcc gcc tat tcc gtc tac ctg aac tca act gag tcc agc act gcc tcg 1764
Ala Ala Tyr Ser Val Tyr Leu Asn Ser Thr Glu Ser Ser Thr Ala Ser
275 280 285

atc ctc ttc ggt gcg gtt gac caa agc aaa tat acc gga agt ctt ggc 1812
Ile Leu Phe Gly Ala Val Asp Gln Ser Lys Tyr Thr Gly Ser Leu Gly
290 295 300

ttg ctt cct atc atc aac acg gct gct tcc tac ggt tac caa aag cct 1860
Leu Leu Pro Ile Ile Asn Thr Ala Ala Ser Tyr Gly Tyr Gln Lys Pro
305 310 315 320

cta agg ctc caa atc acc ctg tct gcc att acg gtc agc gac tcc aga 1908
Leu Arg Leu Gln Ile Thr Leu Ser Ala Ile Thr Val Ser Asp Ser Arg
325 330 335

gga cag caa gca agc att ggt tca gga gct gct gca ctt ctt gat 1956
Gly Gln Gln Ala Ser Ile Gly Ser Gly Ala Ala Ala Leu Leu Asp
340 345 350

# Sequence Listing

```
acc gga acg act ttg acg tat gct cca agc gag att gtc gag aaa ctt        2004
Thr Gly Thr Thr Leu Thr Tyr Ala Pro Ser Glu Ile Val Glu Lys Leu
        355                 360                 365

gct gaa acc cta ggc ttc gac tac agc agc tct gtc ggg gcc tac gtg        2052
Ala Glu Thr Leu Gly Phe Asp Tyr Ser Ser Ser Val Gly Ala Tyr Val
    370                 375                 380

gca aga tgc agg gac gtt gat agc tac gct gtc aac ttc gac ttc cag        2100
Ala Arg Cys Arg Asp Val Asp Ser Tyr Ala Val Asn Phe Asp Phe Gln
385                 390                 395                 400

ggt aaa gtg att gaa gct cct ttg agt tcc ttc ctg att gct ctg caa        2148
Gly Lys Val Ile Glu Ala Pro Leu Ser Ser Phe Leu Ile Ala Leu Gln
                405                 410                 415

acc aac tcc gga gaa gtt tcc tcc tac tgc gca ttg ggt att ttc tcc        2196
Thr Asn Ser Gly Glu Val Ser Ser Tyr Cys Ala Leu Gly Ile Phe Ser
            420                 425                 430

tct gga gac gaa tcc ttc acg ctc ggc gat act ttc ctg cga aac gcc        2244
Ser Gly Asp Glu Ser Phe Thr Leu Gly Asp Thr Phe Leu Arg Asn Ala
        435                 440                 445

tac ttt gtg gct gac ctc gag gga tat caa atc gct ata gct aac gtg        2292
Tyr Phe Val Ala Asp Leu Glu Gly Tyr Gln Ile Ala Ile Ala Asn Val
    450                 455                 460

aac ctg aat cct gga gcc gag caa att gag gtc atc tca ggc aac tcc        2340
Asn Leu Asn Pro Gly Ala Glu Gln Ile Glu Val Ile Ser Gly Asn Ser
465                 470                 475                 480

att cct tct gct tcg tcg gtt tcc gat tac tcc aat acc tgg ggc gcc        2388
Ile Pro Ser Ala Ser Ser Val Ser Asp Tyr Ser Asn Thr Trp Gly Ala
                485                 490                 495

tct gcc acc gct ttg gac act gac agg cct act act ctg gga tct gtg        2436
Ser Ala Thr Ala Leu Asp Thr Asp Arg Pro Thr Thr Leu Gly Ser Val
            500                 505                 510
```

# Sequence Listing

act gct gtg ggc gat gaa aga gtg acc tcg acc aag aag gtt tcg agt 2484
Thr Ala Val Gly Asp Glu Arg Val Thr Ser Thr Lys Lys Val Ser Ser
515 520 525

gtg aag aca agc act tcg tcc ggg tcc ggg tcc act tcg gag tcg tct 2532
Val Lys Thr Ser Thr Ser Ser Gly Ser Gly Ser Thr Ser Glu Ser Ser
530 535 540

acg tcc agt tcg cat tcc agc aat ggc cca agg aca gta ggc ttt agt 2580
Thr Ser Ser Ser His Ser Ser Asn Gly Pro Arg Thr Val Gly Phe Ser
545 550 555 560

ttg tgt gcc gtt ttg tgc gca ttc ttg att tct ata cta gtt gtt tgc 2628
Leu Cys Ala Val Leu Cys Ala Phe Leu Ile Ser Ile Leu Val Val Cys
565 570 575

ta gatctgaagt tctaaggggc tttagtcttc atttatgatt ttttttttatt 2680

tggaccgcct cgaattgttt ttccgacggg tctactttaa agctgcaaga tctcgtttag 2740

cgtcgtttat ttctcgttcg tttagtgaca aaaaaacaga aaaaaaaact ataaaaagcg 2800

gtatataacc tttatatttt gataaacatg agcagcgaaa ttaagctagc accaaaggat 2860

tacgagaagg acaaggagtt cgccaaggct ctgcatggca aggacgccgc gagcgctaca 2920

ggaatgagtg cttgggtgaa gaaggacaag gaagctcaaa aagtcgcgat ggaaggatat 2980

ttcaagcact gggacgggaa aaccgacgag gagactgaaa agtcgagact cgaggactac 3040

tcgacgctca ccaagcacta ctacaacctg gtgacggatt tctacgagta tggatgggga 3100

tcctcgttcc acttttccag atactacaag ggagagccat ttagacaagc t 3151

<210> 2
<211> 576
<212> PRT
<213> Hansenula polymorpha

# Sequence Listing

<400> 2

Met Lys Val Ala Thr Leu Phe Phe Leu Ala Ser Ser Val Cys Val Leu
1 5 10 15

Gly Asp Pro Gln Phe Val Lys Leu Glu Ala Ser Val Leu Arg Gly Ser
20 25 30

Thr Tyr Lys Asp Ser Gln Lys Gly Ala Lys Pro Phe Met Leu Glu Lys
35 40 45

Arg Ala Asp Asp Gly Ser Val Thr Met Glu Leu Gln Asn Ala Gln Ser
50 55 60

Phe Tyr Gln Val Glu Ile Glu Ile Gly Ser Asp Lys Gln Lys Val Gly
65 70 75 80

Val Leu Ile Asp Thr Gly Ser Ser Asp Leu Trp Val Met Asn Ser Asn
85 90 95

Asn Ser Tyr Cys Ser Ser Ser Ser Thr Lys Lys Leu Lys Arg Asp Gly
100 105 110

Pro Ala Asp Ala Leu Gln Lys Gly Arg Asp Leu Ser Asp Leu Tyr Asn
115 120 125

Phe Asn Ser Pro Asn Glu Asp Asn Asn Ala Lys Gly Phe Leu Gly Gly
130 135 140

Trp Gly Asp Leu Thr Thr Val Glu Thr Ala Thr Gln Asp Glu Thr Gln
145 150 155 160

Thr Ala Leu Ala Ala Gln Ala Thr Val Asp Cys Ser Leu Tyr Gly Thr
165 170 175

Phe Asn Pro Ser Thr Ser Asn Ser Phe His Asn Asn Gly Thr Thr Phe
180 185 190

Glu Ile Ser Tyr Ala Asp Arg Thr Phe Ala Arg Gly Thr Trp Gly Tyr
195 200 205

# Sequence Listing

```
Asp Asp Val Thr Phe Asn Gly Val Thr Val Asn Asp Leu Ser Leu Ala
    210                 215                 220

Val Ala Asp Glu Thr Asp Ser Ser Thr Gly Val Phe Gly Ile Gly Leu
225                 230                 235                 240

Arg Glu Leu Glu Thr Thr Tyr Ser Gly Gly Gly Pro Gln His Tyr Ile
                245                 250                 255

Tyr Asp Asn Leu Pro Phe Lys Met Val Asp Gln Gly Leu Ile Asn Arg
            260                 265                 270

Ala Ala Tyr Ser Val Tyr Leu Asn Ser Thr Glu Ser Ser Thr Ala Ser
        275                 280                 285

Ile Leu Phe Gly Ala Val Asp Gln Ser Lys Tyr Thr Gly Ser Leu Gly
    290                 295                 300

Leu Leu Pro Ile Ile Asn Thr Ala Ala Ser Tyr Gly Tyr Gln Lys Pro
305                 310                 315                 320

Leu Arg Leu Gln Ile Thr Leu Ser Ala Ile Thr Val Ser Asp Ser Arg
                325                 330                 335

Gly Gln Gln Ala Ser Ile Gly Ser Gly Ala Ala Ala Ala Leu Leu Asp
            340                 345                 350

Thr Gly Thr Thr Leu Thr Tyr Ala Pro Ser Glu Ile Val Glu Lys Leu
        355                 360                 365

Ala Glu Thr Leu Gly Phe Asp Tyr Ser Ser Ser Val Gly Ala Tyr Val
    370                 375                 380

Ala Arg Cys Arg Asp Val Asp Ser Tyr Ala Val Asn Phe Asp Phe Gln
385                 390                 395                 400

Gly Lys Val Ile Glu Ala Pro Leu Ser Ser Phe Leu Ile Ala Leu Gln
                405                 410                 415
```

# Sequence Listing

Thr Asn Ser Gly Glu Val Ser Ser Tyr Cys Ala Leu Gly Ile Phe Ser
420 425 430

Ser Gly Asp Glu Ser Phe Thr Leu Gly Asp Thr Phe Leu Arg Asn Ala
435 440 445

Tyr Phe Val Ala Asp Leu Glu Gly Tyr Gln Ile Ala Ile Ala Asn Val
450 455 460

Asn Leu Asn Pro Gly Ala Glu Gln Ile Glu Val Ile Ser Gly Asn Ser
465 470 475 480

Ile Pro Ser Ala Ser Ser Val Ser Asp Tyr Ser Asn Thr Trp Gly Ala
485 490 495

Ser Ala Thr Ala Leu Asp Thr Asp Arg Pro Thr Thr Leu Gly Ser Val
500 505 510

Thr Ala Val Gly Asp Glu Arg Val Thr Ser Thr Lys Lys Val Ser Ser
515 520 525

Val Lys Thr Ser Thr Ser Ser Gly Ser Gly Ser Thr Ser Glu Ser Ser
530 535 540

Thr Ser Ser Ser His Ser Ser Asn Gly Pro Arg Thr Val Gly Phe Ser
545 550 555 560

Leu Cys Ala Val Leu Cys Ala Phe Leu Ile Ser Ile Leu Val Val Cys
565 570 575

<210> 3
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

# Sequence Listing

<400> 3
gaagtgcagc agcagctcct gaacc 25

<210> 4
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 4
ggctgatgac ggctcggtca cgatgg 26

<210> 5
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 5
ggacacgcaa gaggtgtctg 20

<210> 6
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

# Sequence Listing

<400> 6
agctcgctac ccggggatcc gcaactttca ttgtgtcaac 40

<210> 7
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 7
gcacatcccc ctttcgccag cctcttcggt gcggttgacc 40

<210> 8
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 8
gctcggctcc aggattcagg 20

<210> 9
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

# Sequence Listing

<400> 9
ggatccccgg gtaccgagct 20

<210> 10
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 10
caccggtagc taatgatccc 20

<210> 11
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 11
cgaacatcca agtgggccga 20

<210> 12
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 12

# Sequence Listing

ctggcgaaag ggggatgtgc 20

<210> 13
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 13
gaattcatga agtgggtaac cttt 24

<210> 14
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 14
taagcctaag gcagcttgac 20

<210> 15
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 15
caagctgcct taggcttatg cagctgctcc ccggtg 36

# Sequence Listing

<210> 16
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 16
actagtgatt tatgggtcct cgatg 25

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR03/01279

## A. CLASSIFICATION OF SUBJECT MATTER

IPC7 C12N 15/31

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7 C12N 15/31

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Blast, PubMed, Delphion "Hansenula", "aspartic protease", "yapsin"

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | WO 98/01535 A1 (Novo Nordisk A/S) 15 January 1998 | 1-14 |
| A | OLSEN, V. ET AL.: "Identification and Characterization of Saccharomyces cerevisiae Yapsin 3 , a New Member of the Yapsin Family of Aspartic Proteases Encoded by the YPS3 Gene", Biochem. J., 1999, Vol.339, pages 407-411 (Biochemical Society) cited in the application | 1-14 |
| A | KANG, H. A. ET AL.: "Efficient Production of Intact Human Parathyroid Hormone in a Saccharomyces cerevisiae Mutant Deficient in Yeast Aspartic Protease 3 (YAP3)" Appl. Microbiol. Biotechnol. 1998, Vol.50, pages 187-192 (Springer-Verlag) cited in the application | 1-14 |

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents,such combination being obvious to a person skilled in the art

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 03 SEPTEMBER 2003 (03.09.2003) | 03 SEPTEMBER 2003 (03.09.2003) |

| Name and mailing address of the ISA/KR | Authorized officer |
|---|---|
| Korean Intellectual Property Office<br>920 Dunsan-dong, Seo-gu, Daejeon 302-701,<br>Republic of Korea<br>Facsimile No. 82-42-472-7140 | KWON, Oh Sig<br><br>Telephone No. 82-42-481-5773 |

Form PCT/ISA/210 (second sheet) (July 1998)

## INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/KR03/01279

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| WO 98/01535 A1 | 15/01/1998 | US6110703 | 29/08/2000 |
| | | JP0513941T | 24/10/2000 |
| | | IL0127583A0 | 28/10/1999 |
| | | ES2187782T3 | 16/06/2003 |
| | | EP0909312B1 | 04/12/2002 |
| | | EP0909312A1 | 21/04/1999 |
| | | DK0909312T3 | 17/03/2003 |
| | | DE69717629C0 | 16/01/2003 |
| | | CN1225126A | 04/08/1999 |
| | | CA2258307AA | 15/01/1998 |
| | | AU3255497A1 | 02/02/1998 |
| | | AT0229068E | 15/12/2002 |